Electromagnetic Wave Theory

JIN AU KONG

VISIT…

ELECTROMAGNETIC WAVE THEORY

JIN AU KONG

Professor of Electrical Engineering
Massachusetts Institute of Technology

A Wiley-Interscience Publication
JOHN WILEY & SONS
New York • Chichester • Brisbane • Toronto • Singapore

1986

Copyright © 1986 by Jin Au Kong

Published by John Wiley & Sons, Inc.

All rights reserved. Published simultaneously in Canada.

Reproduction or translation of any part of this work beyond that permitted by Section 107 or 108 of the 1976 United States Copyright Act without the permission of the copyright owner is unlawful. Requests for permission or further information should be addressed to the Permissions Department, John Wiley & Sons, Inc.

Library of Congress Cataloging in Publication Data

Kong, Jin Au, 1942–
Electromagnetic wave theory.

Includes index.
1. Electromagnetic waves. I. Title.

QC661.K648 1985 530.1'41 85-9554
ISBN 0-471-82823-8

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

PREFACE

This book presents a unified macroscopic theory of electromagnetic waves in accordance with the principle of special relativity from the point of view of the form invariance of Maxwell's equations and the constitutive relations. Topics essential to the understanding of electromagnetic waves are so selected and presented as to make the book a useful graduate text.

Throughout the book electromagnetic waves are our primary concern. For example, when an electromagnetic wave encounters a medium we are concerned with how the wave is affected by the medium, rather than with the reaction of the medium to the wave field. In Chapter I, the fundamental equations and boundary conditions are presented. Time-harmonic fields are studied in Chapter II with the kDB system developed to treat waves in anisotropic and bianisotropic media. Chapter III is devoted to the treatment of reflection, transmission, guidance, and resonance of electromagnetic waves. Starting with the study of Čerenkov radiation, we present antenna theory with simple structures in Chapter IV. Chapter V then elaborates on the various theorems and limiting cases of Maxwell's theory important to the understanding of electromagnetic wave behavior. Scattering by spheres, cylinders, rough surfaces, and volume inhomogeneities are studied in Chapter VI. In Chapter VII, we present Maxwell's theory from the point of view of Lorentz covariance in accordance with the principle of special relativity. The problem section at the end of each chapter provides useful exercise and applications. A solution manual accompanying the book is made available. The various topics in the book can be taught independently, and the material is organized in the order of increasing complexity in mathematical techniques and conceptual abstraction and sophistication.

Great emphasis is placed on the fundamental importance of the

$\overline{k}$ vector in electromagnetic wave theory. The magnitude of the wave vector $\overline{k}$, denoted by k, is called the wavenumber. We shall use subscripts and superscripts for $\overline{k}$ in order to differentiate between different wave vectors and their components instead of defining new symbols. A word about notation is thus in order. A vector $\overline{A}$ is denoted with an overbar, $\hat{A}$ is a unit vector with magnitude equal to unity and denoted with a hat, and dyadic or matrix $\overline{\overline{A}}$ is denoted with a double bar. For coordinate systems, the z axis is always pointing upward, the x axis to the right, and the y axis into the paper. The vector $\overline{A}$ is expressed as $\overline{A} = \hat{x}A_x + \hat{y}A_y + \hat{z}A_z$, with $\hat{x}$, $\hat{y}$, and $\hat{z}$ as the unit vectors in the $\hat{x}$, $\hat{y}$, and $\hat{z}$ directions. Cylindrical coordinates are denoted by ρ, ϕ, z; ϕ is the polar angle with respect to the x axis; $\hat{\rho}$ is the radial unit vector in the $x - y$ plane. Spherical coordinates are denoted by r, θ, ϕ; $\hat{r}$ is the radial unit vector, θ is the angle with respect to the z axis. The real part of a complex quantity A is denoted by A_R, and its imaginary part is denoted by A_I. For the wave vector $\overline{k}$, the coordinate components k are k_x, k_y, k_z in rectangular coordinates; k_ρ, k_ϕ, k_z in cylindrical coordinates; k_r, k_θ, k_ϕ in spherical coordinates. The real and imaginary parts of the z component of $\overline{k}$ in medium 1, for instance, are k_{1zR} and k_{1zI}. For time-harmonic fields we use the time-dependent factor $e^{-i\omega t}$, which leads to familiar equations in quantum theory and facilitates integration in the complex plane. It also leads to the definition of an impedance which is the complex conjugate of that used in circuit theory, but circuit concepts are not emphasized. The field quantities in real space–time, in k space, or in ω space are all distinguished by the use of their explicit independent variables, rather than different symbols or the same symbols with different shapes. For four-dimensional notations, the covariant tensors are denoted with subscript indices and contravariant tensors with superscript indices.

The material in this book has been used in several graduate courses that I have been teaching at the Massachusetts Institute of Technology. The development of the various concepts relies heavily on published work. I have not attempted the task of referring to all relevant publications. The list of books and journal articles in the Reference Section at the end of the book is at best representative and by no means exhaustive. Some of the results contained in the book are taken from many of my research projects, which have been supported by the Joint Service Electronics Program, by grants and contracts from the

National Science Foundation, the National Aeronautics and Space Administration, the Schlumberger-Doll Research, the Office of Naval Research, and the IBM Corporation.

During the writing and preparation of the present version, many people helped. In particular, I would like to acknowledge Soon Yun Poh for his meticulous editing and insightful comments. Mike Tsuk prepared most of the figures with computer graphics and Ruby Li, Sue Wang, and Julie Wu patiently typed and edited several revisions. Over the years, many of my teaching and research assistants provided useful suggestions and proofreading, notably Leung Tsang, Michael Zuniga, Weng Chew, Tarek Habashy, Shun-Lien Chuang, Robert Shin, Jay Kyoon Lee, Apo Sezginer, and Eric Yang. I would like to express my gratitude to them and to the students whose enthusiastic response and feedback give me joy and satisfaction in teaching. Last but not least, I wish to thank my family members, who in many ways made my task of writing this book an enjoyable experience.

J. A. Kong

Cambridge, Massachusetts
September 1985

CONTENTS

Chapter I	**FUNDAMENTAL EQUATIONS**	**1**
1.1	Maxwell's Equations and Constitutive Relations	1
1.2	Wave Equations and Wave Solutions	8
1.3	Conservation Theorems	13
1.4	Polarization	16
1.5	Boundary Conditions	24
	Problems	27
Chapter II	**PROPAGATION IN HOMOGENEOUS MEDIA**	**43**
2.1	Time-Harmonic Fields	44
2.2	Plane Wave Solutions	51
2.3	Plane Waves in Homogeneous Media and the kDB System	60
2.4	Plane Waves in Uniaxial Media	67
2.5	Plane Waves in Gyrotropic Media	73
2.6	Plane Waves in Bianisotropic Media	77
2.7	Plane Waves in Nonlinear Media	79
	Problems	87
Chapter III	**REFLECTION, TRANSMISSION, GUIDANCE, AND RESONANCE**	**103**
3.1	Phase Matching	104
3.2	Reflection and Transmission at a Plane Boundary	110
3.3	Reflection and Transmission by a Layered Medium	120
3.4	Guidance by Conducting Parallel Plates	132
3.5	Guided Waves in Layered Media	150
3.6	Cylindrical Waveguides	164
3.7	Cavity Resonators	182
	Problems	194

Chapter IV	**RADIATION**	**219**
4.1	Čerenkov Radiation	220
4.2	Green's Functions	225
4.3	Hertzian Dipoles	229
4.4	Radiation Fields	237
4.5	Biconical Antennas	247
4.6	Linear Array Antennas	264
4.7	Contour Integration Methods	291
4.8	Integral Formulations for Dipoles in Layered Media	309
4.9	Dipole on One-Layer Medium	319
4.10	Dipole Above Layered Medium	325
	Problems	328
Chapter V	**THEOREMS OF WAVES AND MEDIA**	**357**
5.1	Equivalence Principle	358
5.2	Duality and Complementarity	367
5.3	Mathematical Formulations of the Huygens' Principle	376
5.4	Fresnel and Fraunhofer Diffraction	385
5.5	Reaction and Reciprocity	396
5.6	Stationary Formulas and Rayleigh-Ritz Procedure	405
5.7	Geometrical Optics Limit	415
5.8	Paraxial Limit	436
5.9	Quasi-Static Limits	443
5.10	Quantization of Electromagnetic Waves	448
	Problems	460

Chapter VI	**SCATTERING**	**481**
6.1	Scattering by Spheres	482
6.2	Scattering by a Conducting Cylinder	489
6.3	Scattering by Periodic Rough Surfaces	495
6.4	Scattering by Periodic Media	507
6.5	Scattering by Random Media	517
6.6	Scattering by Random Rough Surfaces	527
6.7	Effective Permittivity for a Volume Scattering Medium	550
	Problems	563
Chapter VII	**ELECTROMAGNETIC THEORY AND SPECIAL RELATIVITY**	**577**
7.1	Lorentz Transformation	578
7.2	Maxwell–Minkowski Theory	581
7.3	Derivation of Transformation Formulas	585
7.4	Transformation of Constitutive Relations	593
7.5	Transformation of Frequency and Wave Vector	599
7.6	Plane Waves in Moving Uniaxial Media	602
7.7	Phase Matching at Moving Boundaries	606
7.8	Guided Waves in a Moving Dielectric Slab	608
7.9	Guided Waves in Moving Gyrotropic Media	611
7.10	Four-Dimensional Notations	614
7.11	Hamilton's Principle and Noether's Theorem	624
	Problems	631
	REFERENCES	**645**
	INDEX	**669**

I

FUNDAMENTAL EQUATIONS

1.1 Maxwell's Equations and Constitutive Relations
a. Constitutive Matrices
b. Isotropic, Anisotropic, and Bianisotropic Media
1.2 Wave Equations and Wave Solutions
1.3 Conservation Theorems
a. Poynting's Theorem
b. Momentum Conservation Theorem
1.4 Polarization
a. Stokes Parameters and Poincaré Sphere
b. Partial Polarization
1.5 Boundary Conditions
a. Stationary Boundaries
b. Moving Boundaries
Problems

1.1 Maxwell's Equations and Constitutive Relations

The fundamental equations of electromagnetic wave theory were established by James Clerk Maxwell in 1873 and experimentally verified by Heinrich Hertz in 1888. Since then electromagnetic wave theory has played a central role in the development of radio, television, optical communications, radar, microwave heating, remote sensing, and numerous other practical applications. Albert Einstein's special theory of relativity in 1905 further asserted the rigorousness and elegance of Maxwell's theory. As a well-established scientific discipline, this sophisticated theoretical structure embodies many principles and concepts which serve as fundamental rules of nature and vital links to other scientific disciplines.

In three-dimensional vector notation, Maxwell's equations are

$$\nabla \times \overline{E}(\bar{r},t) + \frac{\partial}{\partial t}\overline{B}(\bar{r},t) = 0 \tag{1}$$

$$\nabla \times \overline{H}(\bar{r},t) - \frac{\partial}{\partial t}\overline{D}(\bar{r},t) = \overline{J}(\bar{r},t) \tag{2}$$

$$\nabla \cdot \overline{B}(\bar{r},t) = 0 \tag{3}$$

$$\nabla \cdot \overline{D}(\bar{r},t) = \rho(\bar{r},t) \tag{4}$$

where $\overline{E}$, $\overline{B}$, $\overline{H}$, $\overline{D}$, $\overline{J}$, and ρ are real functions of position and time.

$\overline{E}(\bar{r},t)$ = electric field strength	(volts/m)
$\overline{B}(\bar{r},t)$ = magnetic flux density	(webers/m^2)
$\overline{H}(\bar{r},t)$ = magnetic field strength	(amperes/m)
$\overline{D}(\bar{r},t)$ = electric displacement	(coulombs/m^2)
$\overline{J}(\bar{r},t)$ = electric current density	(amperes/m^2)
$\rho(\bar{r},t)$ = electric charge density	(coulombs/m^3)

Equation (1) is Faraday's induction law. Equation (2) is the generalized Ampere's circuit law. Equations (3) and (4) are Gauss' laws for magnetic and electric fields. Taking the divergence of (2) and introducing (4), we find that

$$\nabla \cdot \overline{J}(\bar{r},t) + \frac{\partial}{\partial t}\rho(\bar{r},t) = 0 \tag{5}$$

This is the conservation law for electric charge and current densities. Regarding (5) as a fundamental equation, we see that it can be used to derive (4) by taking the divergence of (2). Equation (3) can also be derived by taking the divergence of (1) which gives $\partial(\nabla \cdot \overline{B}(\bar{r},t))/\partial t = 0$ or that $\nabla \cdot \overline{B}(\bar{r},t)$ is a constant independent of time. Such a constant, if not zero, then implies the existence of magnetic monopoles similar to free electric charges. Since magnetic monopoles have not been found to exist, this constant must be zero and we arrive at (3).

Maxwell's equations are fundamental laws governing the behavior of a electromagnetic field in free space and in media. We have so far made no reference to the various material properties that provide connections to other disciplines of physics, such as plasma physics,

continuum mechanics, solid-state physics, fluid dynamics, statistical physics, thermodynamics, biophysics, etc., all of which interact in one way or another with electromagnetic fields. We did not even mention the Lorentz force law, which constitutes a direct link to mechanics. It is time to state how we are going to account for this vast "outside" world. From the electromagnetic wave point of view, we shall be interested in how electromagnetic fields behave in the presence of media, whether the wave is diffracted, refracted, or scattered. Whatever happens to a medium, whether it is moved or deformed, is of secondary interest. Thus we shall characterize material media by the so-called constitutive relations that can be classified according to the various properties of the media.

The necessity of using constitutive relations to supplement the Maxwell equations is clear from the following mathematical observations. In most problems we shall assume that sources of electromagnetic fields are given. Thus $\overline{J}$ and ρ are known and they satisfy the conservation equation (5). Let us examine the Maxwell equations and see if there are enough equations for the number of unknown quantities. There are a total of 12 scalar unknowns for the four field vectors $\overline{E}$, $\overline{H}$, $\overline{B}$, and $\overline{D}$. As we have learned, (3) and (4) are not independent equations; they can be derived from (1), (2) and (5). The independent equations are (1) and (2), which constitute six scalar equations. Thus we need six more scalar equations. These are the constitutive relations.

a. Constitutive Matrices

Constitutive relations in the most general form can be written as

$$c\overline{D} = \overline{\overline{P}} \cdot \overline{E} + \overline{\overline{L}} \cdot c\overline{B} \tag{6a}$$

$$\overline{H} = \overline{\overline{M}} \cdot \overline{E} + \overline{\overline{Q}} \cdot c\overline{B} \tag{6b}$$

where $c = 3 \times 10^8$ m/s is the velocity of light in vacuum, and $\overline{\overline{P}}$, $\overline{\overline{Q}}$, $\overline{\overline{L}}$, and $\overline{\overline{M}}$ are all 3×3 matrices. Their elements are called *constitutive parameters*. The reason that we write constitutive relations in the present form is based on relativistic considerations. First, the fields $\overline{E}$ and $c\overline{B}$ form a single tensor in four-dimensional space, and so do $c\overline{D}$ and $\overline{H}$. Second, constitutive relations written in the form (6) are Lorentz-covariant. These aspects will be discussed in Chapter VII.

Equation (6) can be rewritten in the form

$$\begin{bmatrix} c\overline{D} \\ \overline{H} \end{bmatrix} = \overline{\overline{C}} \cdot \begin{bmatrix} \overline{E} \\ c\overline{B} \end{bmatrix} \tag{7a}$$

and $\overline{\overline{C}}$ is a 6×6 constitutive matrix:

$$\overline{\overline{C}} = \begin{bmatrix} \overline{\overline{P}} & \overline{\overline{L}} \\ \overline{\overline{M}} & \overline{\overline{Q}} \end{bmatrix} \tag{7b}$$

which has the dimension of admittance.

The constitutive matrix $\overline{\overline{C}}$ may be functions of space–time coordinates, thermodynamical and continuum-mechanical variables, or electromagnetic field strengths. According to the functional dependence of $\overline{\overline{C}}$, we can classify the various media as (i) inhomogeneous if $\overline{\overline{C}}$ is a function of space coordinates, (ii) nonstationary if $\overline{\overline{C}}$ is a function of time, (iii) time-dispersive if $\overline{\overline{C}}$ is a function of time derivatives, (iv) spatial-dispersive if $\overline{\overline{C}}$ is a function of spatial derivatives, (v) nonlinear if $\overline{\overline{C}}$ is a function of the electromagnetic field, and so forth. In the general case $\overline{\overline{C}}$ may be a function of integral-differential operators and coupled to fundamental equations of other physical disciplines.

We have defined constitutive relations by expressing $\overline{D}$ and $\overline{H}$ in terms of $\overline{E}$ and $\overline{B}$. We may also express constitutive relations in the form of $\overline{D}$ and $\overline{B}$ as a function of $\overline{E}$ and $\overline{H}$:

$$\begin{bmatrix} \overline{D} \\ \overline{B} \end{bmatrix} = \overline{\overline{C}}_{EH} \cdot \begin{bmatrix} \overline{E} \\ \overline{H} \end{bmatrix} \tag{8a}$$

where

$$\overline{\overline{C}}_{EH} = \begin{bmatrix} \overline{\overline{\epsilon}} & \overline{\overline{\xi}} \\ \overline{\overline{\zeta}} & \overline{\overline{\mu}} \end{bmatrix} = \frac{1}{c} \begin{bmatrix} \overline{\overline{P}} - \overline{\overline{L}} \cdot \overline{\overline{Q}}^{-1} \cdot \overline{\overline{M}} & \overline{\overline{L}} \cdot \overline{\overline{Q}}^{-1} \\ -\overline{\overline{Q}}^{-1} \cdot \overline{\overline{M}} & \overline{\overline{Q}}^{-1} \end{bmatrix} \tag{8b}$$

Here $\overline{\overline{C}}_{EH}$ is the constitutive matrix under $\overline{E}\,\overline{H}$ representation. To express $\overline{E}$ and $\overline{H}$ in terms of $\overline{B}$ and $\overline{D}$, we write

$$\begin{bmatrix} \overline{E} \\ \overline{H} \end{bmatrix} = \overline{\overline{C}}_{DB} \cdot \begin{bmatrix} \overline{D} \\ \overline{B} \end{bmatrix} \tag{9a}$$

where

$$\overline{\overline{C}}_{DB} = \begin{bmatrix} \overline{\overline{\kappa}} & \overline{\overline{\chi}} \\ \overline{\overline{\gamma}} & \overline{\overline{\nu}} \end{bmatrix} = c \begin{bmatrix} \overline{\overline{P}}^{-1} & -\overline{\overline{P}}^{-1} \cdot \overline{\overline{L}} \\ \overline{\overline{M}} \cdot \overline{\overline{P}}^{-1} & \overline{\overline{Q}} - \overline{\overline{M}} \cdot \overline{\overline{P}}^{-1} \cdot \overline{\overline{L}} \end{bmatrix} \tag{9b}$$

Here $\overline{\overline{C}}_{DB}$ is the constitutive matrix under $\overline{D}\,\overline{B}$ representation. The other possible construction for expressing $\overline{E}$ and $\overline{B}$ in terms of $\overline{H}$ and $\overline{D}$ is not shown because it will not be needed in later developments.

b. Isotropic, Anisotropic, and Bianisotropic Media

In the definition of the constitutive relations, the constitutive matrices $\overline{\overline{L}}$ and $\overline{\overline{M}}$ relate electric and magnetic fields. When $\overline{\overline{L}}$ and $\overline{\overline{M}}$ are not identically zero, the medium is *bianisotropic*. When there is no coupling between electric and magnetic fields, $\overline{\overline{L}} = \overline{\overline{M}} = 0$ and the medium is *anisotropic*. For an anisotropic medium, if $\overline{\overline{P}} = c\epsilon\overline{\overline{I}}$ and $\overline{\overline{Q}} = (1/c\mu)\overline{\overline{I}}$ with $\overline{\overline{I}}$ denoting the 3×3 unit matrix, the medium is *isotropic*.

Isotropic Media. The constitutive relations for an isotropic medium can be written simply as

$$\overline{D} = \epsilon\overline{E} \qquad \text{where } \epsilon = \text{permittivity} \tag{10a}$$

$$\overline{B} = \mu\overline{H} \qquad \text{where } \mu = \text{permeability} \tag{10b}$$

Thus the field vector $\overline{E}$ is parallel to $\overline{D}$ and the field vector $\overline{H}$ is parallel to $\overline{B}$. In free space void of any matter, $\mu = \mu_o$ and $\epsilon = \epsilon_o$,

$$\mu_o = 4\pi \times 10^{-7} \qquad \text{henry/meter}$$

$$\epsilon_o \approx 8.85 \times 10^{-12} \qquad \text{farad/meter}$$

Inside a material medium, the permittivity ϵ is determined by the electrical properties of the medium and the permeability μ by the magnetic properties of the medium.

A dielectric material can be described by a free-space part and a part that is due to the material alone. The material part can be characterized by a polarization vector $\overline{P}$ such that $\overline{D} = \epsilon_o\overline{E} + \overline{P}$. The polarization $\overline{P}$ symbolizes the electric dipole moment per unit volume of the dielectric material. In the presence of an external electric

field, the polarization vector may be caused by induced dipole moments, alignment of the permanent dipole moments of the medium, or migration of ionic charges.

A magnetic material can also be described by a free-space part and a part characterized by a magnetization vector $\overline{M}$ such that $\overline{B} = \mu_o \overline{H} + \mu_o \overline{M}$. A medium is diamagnetic if $\mu \leq \mu_o$, paramagnetic if $\mu \geq \mu_o$. Diamagnetism is caused by induced magnetic moments that tend to oppose the externally applied magnetic field. Paramagnetism is due to alignment of magnetic moments. When placed in an inhomogeneous magnetic field, a diamagnetic material tends to move toward regions of weaker magnetic field, and a paramagnetic material toward regions of stronger magnetic field. Ferromagnetism and antiferromagnetism are highly nonlinear effects. Ferromagnetic substances are characterized by spontaneous magnetization below the Curie temperature. The medium also depends on the history of applied fields, and in many instances the magnetization curve forms a hysteresis loop. In an antiferromagnetic material, the spins form sub-lattices that become spontaneously magnetized in an antiparallel arrangement below the Néel temperature.

Anisotropic Media. The constitutive relations for anisotropic media are usually written in the $\overline{E}\,\overline{H}$ representation as

$$\overline{D} = \overline{\overline{\epsilon}} \cdot \overline{E} \qquad \text{where } \overline{\epsilon} = \text{permittivity tensor} \tag{11a}$$

$$\overline{B} = \overline{\overline{\mu}} \cdot \overline{H} \qquad \text{where } \overline{\mu} = \text{permeability tensor} \tag{11b}$$

The field vector $\overline{E}$ is no longer parallel to $\overline{D}$, and the field vector $\overline{H}$ is no longer parallel to $\overline{B}$. A medium is *electrically anisotropic* if it is described by the permittivity tensor $\overline{\overline{\epsilon}}$ and *magnetically anisotropic* if it is described by the permeability tensor $\overline{\overline{\mu}}$. Note that a medium can be both electrically and magnetically anisotropic.

Crystals are described in general by symmetric permittivity tensors. There always exists a coordinate transformation that transforms a symmetric matrix into a diagonal matrix. In this coordinate system, called the *principal system,*

$$\overline{\overline{\epsilon}} = \begin{bmatrix} \epsilon_x & 0 & 0 \\ 0 & \epsilon_y & 0 \\ 0 & 0 & \epsilon_z \end{bmatrix} \tag{12}$$

The three coordinate axes are referred to as the principal axes of the crystal. For cubic crystals, $\epsilon_x = \epsilon_y = \epsilon_z$ and they are isotropic. In

tetragonal, hexagonal, and rhombohedral crystals, two of the three parameters are equal. Such crystals are *uniaxial.* Here there is a two-dimensional degeneracy; the principal axis that exhibits this anisotropy is called the *optic axis.* For a uniaxial crystal with

$$\bar{\bar{\epsilon}} = \begin{bmatrix} \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon_z \end{bmatrix} \tag{13}$$

the z axis is the optic axis. The crystal is *positive uniaxial* if $\epsilon_z > \epsilon$; it is *negative uniaxial* if $\epsilon_z < \epsilon$. In orthorhombic, monoclinic, and triclinic crystals, all three crystallographic axes are unequal. We have $\epsilon_x \neq \epsilon_y \neq \epsilon_z$, and the medium is *biaxial.*

Bianisotropic Media. For isotropic or anisotropic media, the constitutive relations relate the two electric field vectors and the two magnetic field vectors by either a scalar or a tensor. Such media become polarized when placed in an electric field and become magnetized when placed in a magnetic field. A bianisotropic medium provides the cross coupling between the electric and magnetic fields. When placed in an electric or a magnetic field, a bianisotropic medium becomes both polarized and magnetized.

Magnetoelectric materials, theoretically predicted by Dzyaloshinskii [1959] and by Landau and Lifshitz [1957], were observed experimentally in 1960 by Astrov in antiferromagnetic chromium oxide. The constitutive relations that Dzyaloshinskii proposed for chromium oxide have the following form:

$$\overline{D} = \begin{bmatrix} \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon_z \end{bmatrix} \cdot \overline{E} + \begin{bmatrix} \xi & 0 & 0 \\ 0 & \xi & 0 \\ 0 & 0 & \xi_z \end{bmatrix} \cdot \overline{H} \tag{14a}$$

$$\overline{B} = \begin{bmatrix} \xi & 0 & 0 \\ 0 & \xi & 0 \\ 0 & 0 & \xi_z \end{bmatrix} \cdot \overline{E} + \begin{bmatrix} \mu & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu_z \end{bmatrix} \cdot \overline{H} \tag{14b}$$

It was then shown by Indenbom [1960] and by Birss [1963] that 58 magnetic crystal classes can exhibit the magnetoelectric effect. Rado [1964] proved that the effect is not restricted to antiferromagnetics; ferromagnetic gallium iron oxide is also magnetoelectric.

In 1948, the gyrator was introduced by Tellegen as a new element, in addition to the resistor, the capacitor, the inductor, and the ideal

transformer, for describing a network. To realize his new network element, Tellegen conceived of a medium possessing constitutive relations of the form

$$\overline{D} = \epsilon \overline{E} + \xi \overline{H} \tag{15a}$$

$$\overline{B} = \xi \overline{E} + \mu \overline{H} \tag{15b}$$

where $\xi^2/\mu\epsilon$ is nearly equal to 1. Tellegen considered that the model of the medium had elements possessing permanent electric and magnetic dipoles parallel or antiparallel to each other, so that an applied electric field that aligns the electric dipoles simultaneously aligns the magnetic dipoles; and a magnetic field that aligns the magnetic dipoles simultaneously aligns the electric dipoles. Tellegen also wrote general constitutive relations (8) and examined the symmetry properties by energy conservation.

Media in motion were the first bianisotropic media to receive attention in electromagnetic theory. In 1888, Roentgen discovered that a moving dielectric becomes magnetized when it is placed in an electric field. In 1905, Wilson showed that a moving dielectric in a uniform magnetic field becomes electrically polarized. Almost any medium becomes bianisotropic when it is in motion. In Chapter VII, we shall derive constitutive relations for uniformly moving media using the Lorentz transformation of field vectors.

The bianisotropic description of material has fundamental importance from the point of view of relativity. The principle of relativity postulates that all physical laws of nature must be characterized by mathematical equations that are form-invariant from one observer to the other. For electromagnetic theory, the Maxwell equations are form-invariant with respect to all observers, although the numerical values of the field quantities may vary from one observer to another. The constitutive relations are form-invariant when they are written in bianisotropic form. In Chapter VII we shall treat special relativity and electromagnetic theory in detail.

1.2 Wave Equations and Wave Solutions

The Maxwell equations in differential form are valid at all times for every point in space. First we shall investigate solutions to the Maxwell equations in regions void of source, namely in regions where $\overline{J} = \rho = 0$.

This of course does not mean that there is no source anywhere in all space. Sources must exist outside the regions of interest in order to produce fields in these regions. Thus in source-free homogeneous isotropic media, Maxwell's equations become

$$\nabla \times \overline{E} = -\mu \frac{\partial}{\partial t} \overline{H} \tag{1}$$

$$\nabla \times \overline{H} = \epsilon \frac{\partial}{\partial t} \overline{E} \tag{2}$$

$$\nabla \cdot \overline{H} = 0 \tag{3}$$

$$\nabla \cdot \overline{E} = 0 \tag{4}$$

We shall consider free space with permittivity $\epsilon = \epsilon_o$ and permeability $\mu = \mu_o$. A wave equation can be easily derived by taking the curl of (1) and substituting (2). We find, upon using the vector identity $\nabla \times (\nabla \times \overline{E}) = \nabla (\nabla \cdot \overline{E}) - \nabla^2 \overline{E}$ and (4),

$$\nabla^2 \overline{E} - \mu\epsilon \frac{\partial^2}{\partial t^2} \overline{E} = 0 \tag{5}$$

The Laplacian operator ∇^2 in a rectangular coordinate system is

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

The simplest solution to (5) is

$$\overline{E} = \hat{x} E_0 \cos(kz - \omega t) = \hat{x} E_x(z,t) \tag{6}$$

Substituting (6) in (5) we find that the following equation, called the dispersion relation, which relates ω and k must be satisfied:

$$k^2 = \omega^2 \mu\epsilon \tag{7}$$

There are two points of view useful in the study of a space–time varying quantity such as $E_x(z,t)$. The first is to examine the time variation at fixed points in space. The second is to examine spatial variation at fixed times, a process that amounts to taking a series of pictures.

We first fix our attention to one particular point in space, say $z = 0$. We then have the electric vector $E_x(z,t) = E_0 \cos \omega t$. Plotted as a function of time in Figure 1.2.1, we find that the waveform repeats

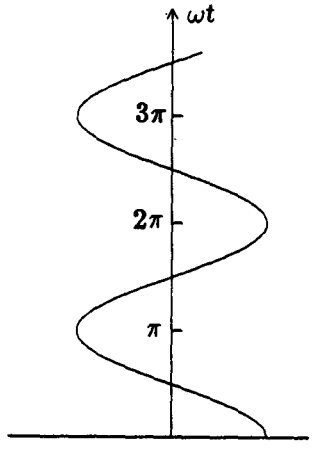


Figure 1.2.1 Electric field magnitude as a function of time at $z = 0$.

itself in time as $\omega t = 2m\pi$ for any integer m. The period is defined as the time T for which $\omega T = 2\pi$. The frequency f is defined as $f = 1/T$ which gives

$$f = \frac{\omega}{2\pi} \tag{8}$$

Since $\omega = 2\pi f$, ω is the angular frequency of the wave. Often we simply refer to frequency ω, as ω is more commonly encountered than f.

To examine wave behavior from the other point of view, we let $\omega t = 0$ and plot $E_x(z,t)$ in Figure 1.2.2a. The waveform repeats itself in space when $kz = 2m\pi$ for integer values of m. The wavelength λ is defined as the distance for which $k\lambda = 2\pi$. Thus $\lambda = 2\pi/k$, or

$$k = \frac{2\pi}{\lambda} \tag{9}$$

We call k the wavenumber which is equal to the number of wavelengths in a distance of 2π and has the dimension inverse length.

In Figures 1.2.2b and 1.2.2c we plot $E_x(z,t)$ at two progressive times $\omega t = \pi/2$ and $\omega t = \pi$. We observe that a constant phase front represented by A with $kz - \omega t = \textit{constant}$ moves along the z axis with a velocity

$$v = \frac{dz}{dt} = \frac{\omega}{k} \tag{10}$$

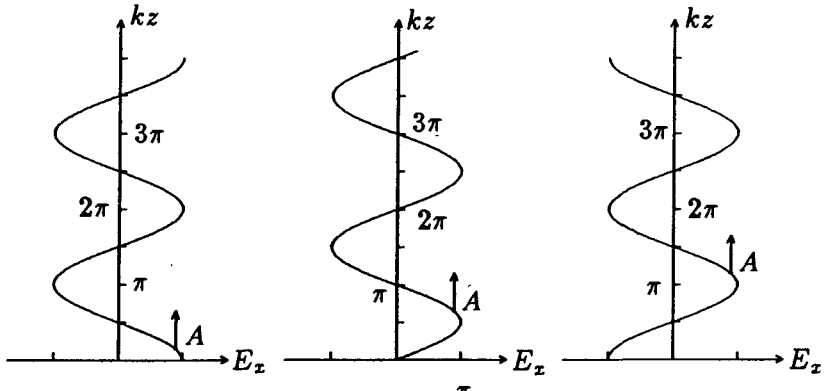


Figure 1.2.2 Electric field magnitude vs. kz at different times.

By virtue of the dispersion relation in (7), we see that $v = (\mu\epsilon)^{-1/2} = (\mu_o\epsilon_o)^{-1/2}$, which is equal to the velocity of light in free space c.

In this book we shall place great emphasis on the use of k, as we shall learn that the wavenumber concept is, in fact, more fundamental in electromagnetic wave theory than both of the more popular concepts of wavelength λ and frequency f. In Figure 1.2.3, we illustrate the electromagnetic wave spectrum according to the free space wavenumber $k = k_o$. The corresponding values of frequency and wavelength are $f = ck/2\pi$ and $\lambda = 2\pi/k$. The photon energy in electron-volts is calculated from $\hbar\omega = \hbar ck$ where $\hbar = 1.05 \times 10^{-34}$ Joule-sec is Planck's constant divided by 2π and the electron charge is 1.6×10^{-19} coulombs.

The magnetic field vector $\overline{H}$ for the wave solution can be found from either (1) or (2), which gives

$$\overline{H}(z,t) = \hat{y}\sqrt{\frac{\epsilon}{\mu}}\, E_0 \cos(kz - \omega t) \tag{11}$$

We see that the magnetic field vector is perpendicular to the electric field vector and that both are perpendicular to the direction of propagation of the wave.

The following is a solution for the electric field $\overline{E}(\bar{r},t)$ to the wave

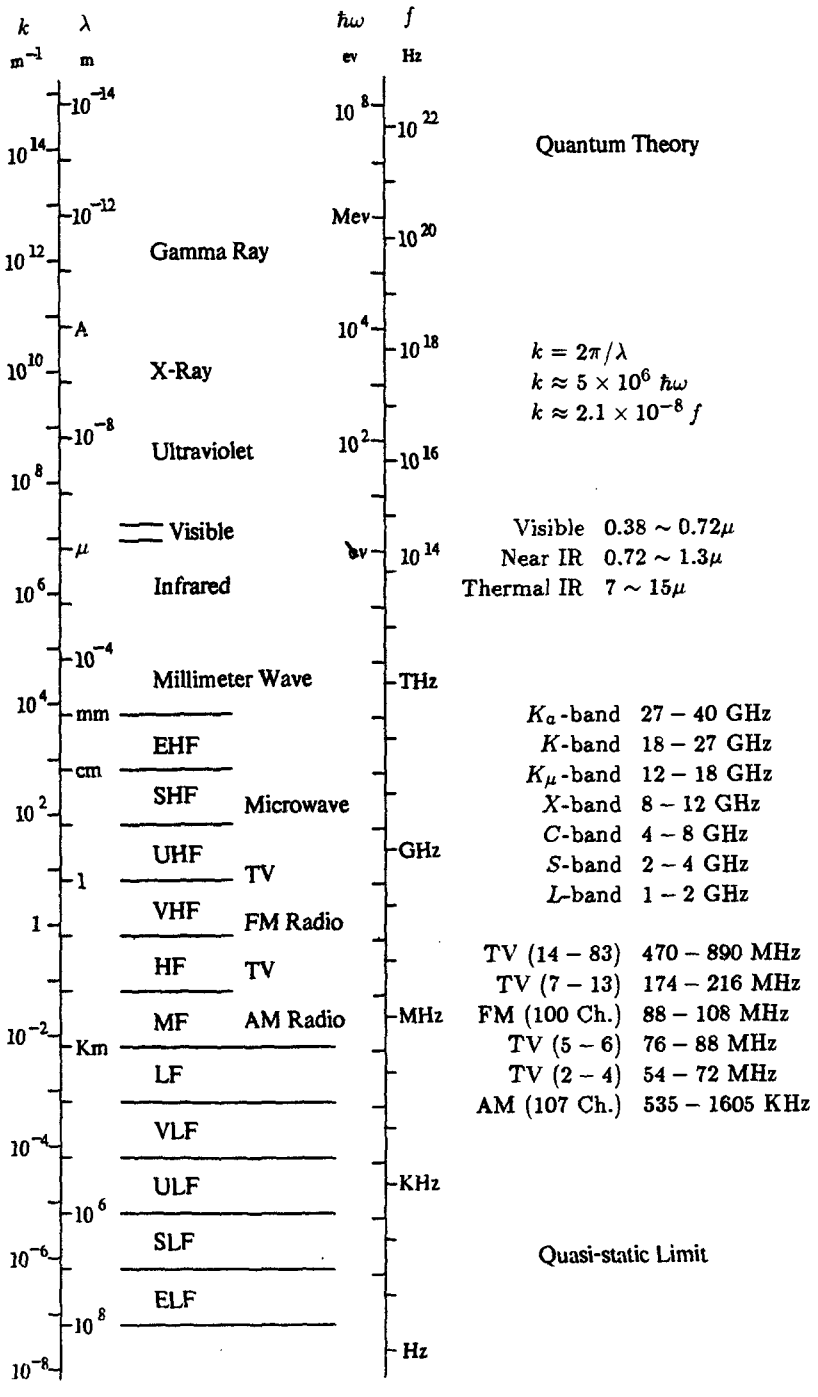


Figure 1.2.3 Electromagnetic Wave Spectrum.

equation (5)

$$\overline{E}(\overline{r},t) = \overline{E}_0 \cos(k_x x + k_y y + k_z z - \omega t) \tag{12}$$

where $\overline{E}_0$ is a constant vector. We define a wave vector

$$\overline{k} = \hat{x}k_x + \hat{y}k_y + \hat{z}k_z$$

and a position vector

$$\overline{r} = \hat{x}x + \hat{y}y + \hat{z}z$$

The field solution then takes the form

$$\overline{E}(\overline{r},t) = \overline{E}_0 \cos(\overline{k}\cdot\overline{r} - \omega t)$$

The wave vector $\overline{k}$ is often referred to simply as the $\overline{k}$ vector. At any instant the field is a constant when $\overline{k}\cdot\overline{r}$ is a constant. The equation $\overline{k}\cdot\overline{r} = \textit{constant}$ defines a plane.

The magnetic field $\overline{H}$ has the same space–time dependence as $\overline{E}$. Substituting (11) and (12) into the Maxwell equations, we find

$$\begin{aligned}
\overline{k}\times\overline{E} &= \omega\mu\overline{H} \\
\overline{k}\times\overline{H} &= -\omega\epsilon\overline{E} \\
\overline{k}\cdot\overline{H} &= 0 \\
\overline{k}\cdot\overline{E} &= 0
\end{aligned}$$

Thus $\overline{E}$ and $\overline{H}$ are both mutually perpendicular and perpendicular to the wave vector $\overline{k}$.

1.3 Conservation Theorems

a. Poynting's Theorem

Energy conservation immediately follows from Maxwell's equations. Dot-multiply Faraday's law (1.1.1) by $\overline{H}$, Ampere's law (1.1.2) by $\overline{E}$ and subtract. By making use of the vector identity $\nabla\cdot(\overline{E}\times\overline{H}) = \overline{H}\cdot\nabla\times\overline{E} - \overline{E}\cdot\nabla\times\overline{H}$, we obtain Poynting's theorem

$$\nabla\cdot(\overline{E}\times\overline{H}) + \overline{H}\cdot\frac{\partial\overline{B}}{\partial t} + \overline{E}\cdot\frac{\partial\overline{D}}{\partial t} = -\overline{E}\cdot\overline{J} \tag{1}$$

The Poynting vector

$$\overline{S} = \overline{E} \times \overline{H} \tag{2}$$

is interpreted as the power flow density with dimension of watts/m^2, and $\overline{H} \cdot (\partial \overline{B}/\partial t) + \overline{E} \cdot (\partial \overline{D}/\partial t)$ represents the time rate of change of the stored electric and magnetic energy. On the right-hand side of (1), $-\overline{E} \cdot \overline{J}$ is the power supplied by the current $\overline{J}$.

As an example consider the simple wave solution

$$\overline{E} = \hat{x} E_0 \cos(kz - \omega t) \tag{3a}$$

$$\overline{H} = \hat{y} \sqrt{\frac{\epsilon}{\mu}} E_0 \cos(kz - \omega t) \tag{3b}$$

The Poynting vector is calculated to be

$$\overline{S} = \hat{z} \sqrt{\frac{\epsilon}{\mu}} E_0^2 \cos^2(kz - \omega t) \tag{4}$$

The time-average Poynting's vector power density is given by

$$< \overline{S} >= \frac{1}{T} \int_0^T dt \, \overline{S} = \hat{z} \frac{E_0^2}{2\eta} \tag{5}$$

where $\eta = (\mu/\epsilon)^{1/2}$ is the characteristic impedance.

For isotropic media, we find

$$\overline{H} \cdot \frac{\partial}{\partial t}(\mu \overline{H}) = \frac{\partial}{\partial t} \left[\frac{1}{2} \mu \overline{H} \cdot \overline{H} \right]$$

and

$$\overline{E} \cdot \frac{\partial}{\partial t}(\epsilon \overline{E}) = \frac{\partial}{\partial t} \left[\frac{1}{2} \epsilon \overline{E} \cdot \overline{E} \right]$$

In the source-free region we also have $\overline{J} = 0$. Poynting's theorem thus becomes

$$\nabla \cdot (\overline{E} \times \overline{H}) + \frac{\partial}{\partial t}(W_e + W_m) = 0 \tag{6}$$

where

$$W_e = \frac{1}{2} \epsilon |\overline{E}|^2 \tag{7}$$

is the stored electric energy density and

$$W_m = \frac{1}{2}\mu|\overline{H}|^2 \tag{8}$$

is the stored magnetic energy density. Substituting (3)–(5) in (6)–(8) we see clearly that Poynting's theorem is satisfied.

We can show that for linear bianisotropic media with symmetric constitutive matrix $\overline{\overline{C}}_{EH}$, Poynting's theorem becomes

$$\frac{\partial W}{\partial t} + \nabla \cdot \overline{S} = -p \tag{9}$$

where

$$W = \frac{1}{2}(\overline{D}\cdot\overline{E} + \overline{B}\cdot\overline{H}) = \text{total stored energy} \tag{10}$$

$$\overline{S} = \overline{E}\times\overline{H} = \text{Poynting's power density vector} \tag{11}$$

$$-p = -\overline{J}\cdot\overline{E} = \text{power supplied by external agent} \tag{12}$$

This is a general form expressing conservation of energy.

b. Momentum Conservation Theorem

The other fundamental law that relates electromagnetism to mechanics is the Lorentz force law

$$\overline{f} = \rho\overline{E} + \overline{J}\times\overline{B} \tag{13}$$

It can be used to define the fields $\overline{E}$ and $\overline{B}$. Substituting Maxwell's equations for ρ and $\overline{J}$ in the right-hand side, we find that, for linear, homogeneous bianisotropic media with symmetric constitutive matrix $\overline{\overline{C}}_{EH}$,

$$\overline{f} = -\frac{\partial}{\partial t}(\overline{D}\times\overline{B}) - \nabla\cdot\left[\frac{1}{2}(\overline{D}\cdot\overline{E} + \overline{B}\cdot\overline{H})\overline{\overline{I}} - \overline{D}\,\overline{E} - \overline{B}\,\overline{H}\right] \tag{14}$$

where $\overline{\overline{I}}$ is a unit dyad with diagonal elements equal to 1 and all off-diagonal elements equal to zero.

The interpretation of the terms is

$$\overline{G} = \overline{D}\times\overline{B} = \text{momentum density vector} \tag{15}$$

$$\overline{\overline{T}} = \frac{1}{2}(\overline{D}\cdot\overline{E} + \overline{B}\cdot\overline{H})\overline{\overline{I}} - \overline{D}\,\overline{E} - \overline{B}\,\overline{H}$$
$$= \text{Maxwell stress tensor} \tag{16}$$

Thus we have the theorem

$$\frac{\partial \overline{G}}{\partial t} + \nabla \cdot \overline{\overline{T}} = -\overline{f} \tag{17}$$

which expresses conservation of momentum. This is in a form similar to Poynting's theorem (9) except that it is now a vector equation.

1.4 Polarization

The polarization of a wave is conventionally defined by the time variation of the tip of the electric field vector $\overline{E}$ at a fixed point in space. If the tip moves along a straight line, the wave is linearly polarized. When the locus of the tip is a circle, the wave is circularly polarized. For an elliptically polarized wave, the tip of $\overline{E}$ describes an ellipse. If the right-hand thumb points in the direction of propagation while the fingers point in the direction of the tip motion, the wave is defined as right-hand polarized. The wave is left-hand polarized when it is described by the left-hand thumb and fingers.

To facilitate a mathematical discussion of polarization, we decompose the $\overline{E}$ vector of a wave into two components perpendicular to the $\overline{k}$ vector. For a specific point in space, we write

$$\overline{E}(t) = \hat{h}E_h + \hat{v}E_v = \hat{h}e_h \cos(\omega t - \psi_h) + \hat{v}e_v \cos(\omega t - \psi_v) \tag{1}$$

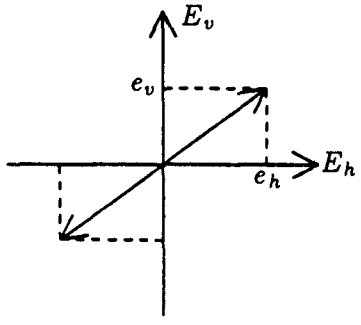
where $\overline{k}$, $\hat{h}$ and $\hat{v}$ form an orthogonal system mutually perpendicular to one another. The locus of the tip $\overline{E}(t)$ is determined by eliminating the time t using the two components E_h and E_v.

Linear Polarization

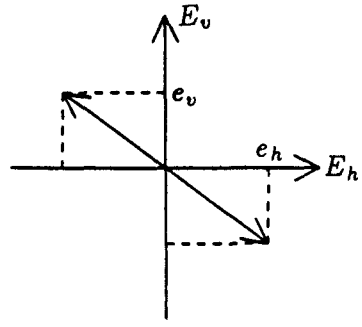
When ψ_h and ψ_v differ by an integer number of 2π, the two components are in phase. The wave is linearly polarized and the straight-line locus traverses the first and third quadrants [Fig. 1.4.1a]. When ψ_h and ψ_v differ by an odd integer multiple of π, the two components are 180° out of phase, and the straight-line locus traverses the second and fourth quadrants [Fig. 1.4.1b].

Circular Polarization

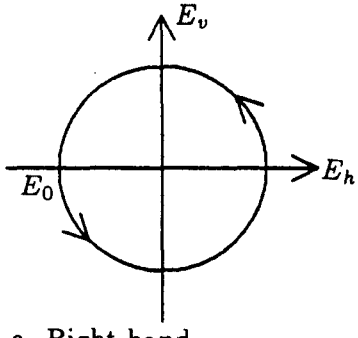
The magnitudes of the two components are equal, $e_h = e_v = e_0$, and the phases differ by 90°. Consider the case $\psi_v - \psi_h = \pi/2$. We



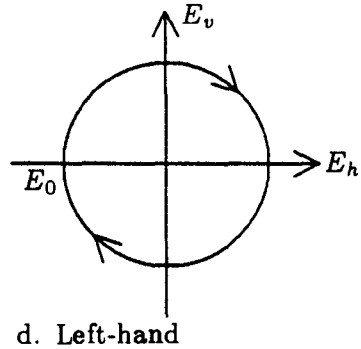
a. Linear polarization



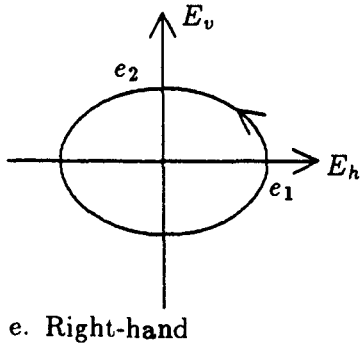
b. Linear polarization



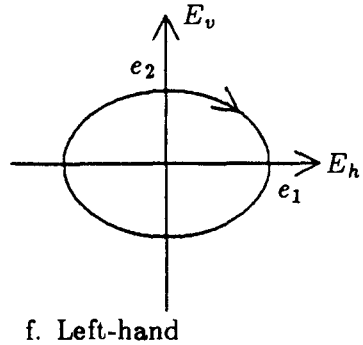
c. Right-hand
circular polarization



d. Left-hand
circular polarization



e. Right-hand
elliptical polarization



f. Left-hand
elliptical polarization

Figure 1.4.1 Polarizations.

have

$$\overline{E}(t) = e_0 [\hat{h} \cos(\omega t - \psi_h) + \hat{v} \sin(\omega t - \psi_h)] \tag{2}$$

It can be seen that while the h component is at its maximum of e_0, the v component is zero. As time progresses, the v component increases and the h component decreases. The tip of $\overline{E}$ rotates from the positive E_h axis to the positive E_v axis [Fig. 1.4.1c]. Elimination of t from the h and v components in (2) yields a circle of radius E_0, $E_h^2 + E_v^2 = e_0^2$. Thus the wave is right-hand circularly polarized. Similar reasoning shows that with $\psi_v - \psi_h = -\pi/2$, the wave is left-hand circularly polarized [Fig. 1.4.1d].

Elliptical Polarization

In general, a polarized wave has elliptical polarization; that is, when time is eliminated from the two components of $\overline{E}$, the resultant equation describes an ellipse.

First consider the case when $\psi_h = \psi_v = \psi_0$

$$\overline{E}(t) = \hat{h} e_1 \cos(\omega t - \psi_0) \pm \hat{v} e_2 \sin(\omega t - \psi_0) = \hat{h} E_1 + \hat{v} E_2 \tag{3}$$

Elimination of time yields an ellipse

$$\left(\frac{E_h}{e_1}\right)^2 + \left(\frac{E_v}{e_2}\right)^2 = 1$$

With the plus sign in (3), we have a right-hand elliptically polarized wave [Fig. 1.4.1e] and with the minus sign a left-hand elliptically polarized wave [Fig. 1.4.1f].

This discussion can be summarized in Figure 1.4.2 where we plot the magnitude ratio

$$E_v/E_h = A \tag{4}$$

and the phase difference

$$\psi = \psi_v - \psi_h \tag{5}$$

on a plane. If $\psi = m\pi$, the wave is linearly polarized. If $A = 1$ and $\psi = \pi/2$, the wave is right-hand circularly polarized. For $A = 1$ and $\psi = -\pi/2$, the wave is left-hand circularly polarized. Otherwise the wave is elliptically polarized. The polarization is right-handed if the phase difference is between zero and π, and left-handed if ψ is between π and 2π.

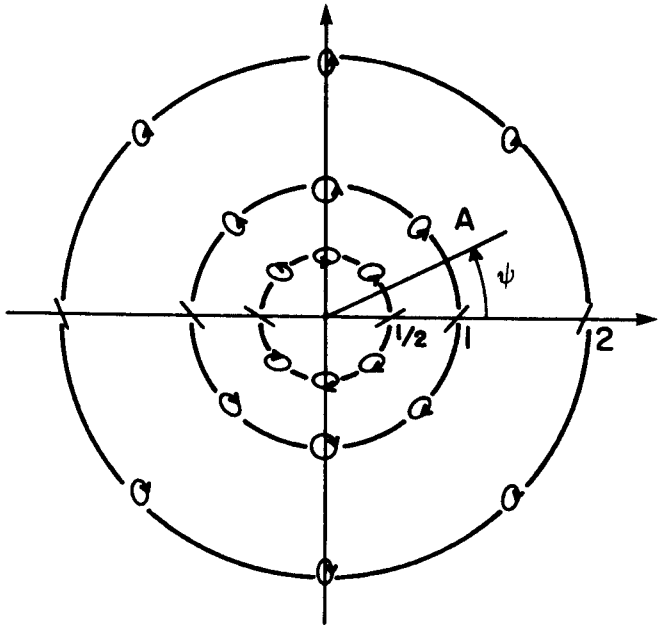


Figure 1.4.2 Polarizations for various values of ψ and A.

a. Stokes Parameters and Poincaré Sphere

The general polarization states are more popularly described with the Poincaré sphere as discussed below. Consider the elliptical polarization as given by (1), from which we find

$$\frac{E_h}{e_h}\sin\psi_v - \frac{E_v}{e_v}\sin\psi_h = \cos\omega t\sin\psi \tag{6a}$$

$$\frac{E_h}{e_h}\cos\psi_v - \frac{E_v}{e_v}\cos\psi_h = -\sin\omega t\sin\psi \tag{6b}$$

Eliminating the time dependence t, we obtain the equation

$$\left(\frac{E_h}{e_h}\right)^2 + \left(\frac{E_v}{e_v}\right)^2 - 2\frac{E_hE_v}{e_he_v}\cos\psi = \sin^2\psi \tag{7}$$

The polarization ellipse is plotted in Figure 1.4.3 where the major axis makes the angle α with the E_h axis. The shape and orientation of the ellipse can be specified, with e_1 denoting the major axis and e_2 the minor axis, by the angle β,

$$\tan\beta = \pm\frac{e_2}{e_1} \tag{8}$$

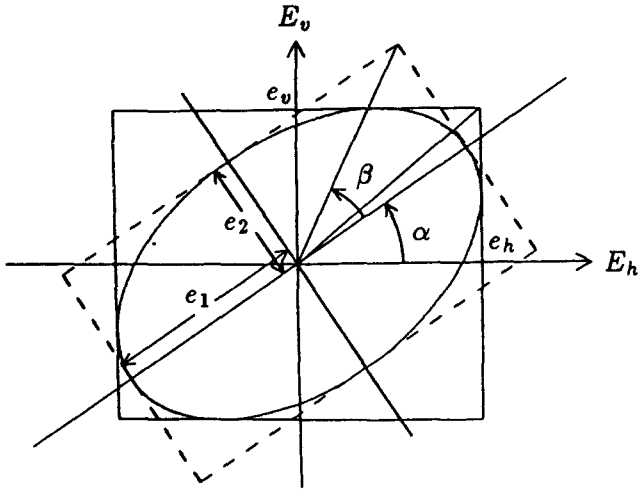


Figure 1.4.3 Elliptical polarization.

where for right-hand polarization $0 \leq \beta \leq \pi/4$ and for left-hand polarization $-\pi/4 \leq \beta \leq 0$. In view of (3) and (8), we have from coordinate transformation

$$e_1 \cos(\omega t - \psi_0) = E_1 = E_h \cos\alpha + E_v \sin\alpha \tag{9a}$$

$$e_1 \tan\beta \sin(\omega t - \psi_0) = E_2 = -E_h \sin\alpha + E_v \cos\alpha \tag{9b}$$

where $0 \leq \alpha \leq \pi$. Substituting the components E_h and E_v of (1) in (9) and comparing the coefficients of $\cos\omega t$ and $\sin\omega t$, we obtain

$$e_1 \cos\psi_0 = e_h \cos\psi_h \cos\alpha + e_v \cos\psi_v \sin\alpha \tag{10a}$$

$$e_1 \sin\psi_0 = e_h \sin\psi_h \cos\alpha + e_v \sin\psi_v \sin\alpha \tag{10b}$$

$$e_1 \tan\beta \cos\psi_0 = -e_h \sin\psi_h \sin\alpha + e_v \sin\psi_v \cos\alpha \tag{10c}$$

$$e_1 \tan\beta \sin\psi_0 = e_h \cos\psi_h \sin\alpha - e_v \cos\psi_v \cos\alpha \tag{10d}$$

Eliminating ψ_0 from (10a) and (10b) by squaring and adding, we find

$$e_1^2 = e_h^2 \cos^2\alpha + e_v^2 \sin^2\alpha + e_h e_v \sin 2\alpha \cos\psi \tag{11a}$$

Similarly from (10c) and (10d), we have

$$e_1^2 \tan^2\beta = e_2^2 = e_h^2 \sin^2\alpha + e_v^2 \cos^2\alpha - e_h e_v \sin 2\alpha \cos\psi \tag{11b}$$

Note that the sum of (11a) and (11b) gives $e_1^2 + e_2^2 = e_h^2 + e_v^2$.

Multiplying (10a) by (10c), (10b) by (10d) and adding, we again eliminate ψ_0 and obtain

$$e_1^2 \tan\beta = e_h e_v \sin\psi \tag{11c}$$

Finally we multiply (10a) by (10d) and subtract from (10b) multiplied by (10c), which yield

$$2e_h e_v \cos\psi = (e_h^2 - e_v^2)\tan 2\alpha \tag{11d}$$

Equation (11) will be used in the following discussion on Stokes parameters and the Poincaré sphere.

To facilitate the discussion of polarization states of electromagnetic waves, the four Stokes parameters pertaining to $\overline{E}(t)$ given in (1) are defined as follows :

$$I = \frac{1}{\eta}(e_h^2 + e_v^2) \tag{12a}$$

$$Q \triangleq \frac{1}{\eta}(e_h^2 - e_v^2) \tag{12b}$$

$$U = \frac{2}{\eta} e_h e_v \cos\psi \tag{12c}$$

$$V = \frac{2}{\eta} e_h e_v \sin\psi \tag{12d}$$

Notice that $I^2 = Q^2 + U^2 + V^2$.

Adding (11a) and (11b) yields

$$e_1^2 = \eta I \cos^2\beta \tag{13}$$

Subtracting (11b) from (11a) and making use of (11d) and (13), we find

$$Q = \frac{1}{\eta}(e_h^2 - e_v^2) = I\cos 2\alpha \cos 2\beta \tag{14a}$$

In terms of I, we find from (12c), (11d) and (14a)

$$U = I \sin 2\alpha \cos 2\beta \tag{14b}$$

and from (12d), (11c) and (13)

$$V = I \sin 2\beta \tag{14c}$$

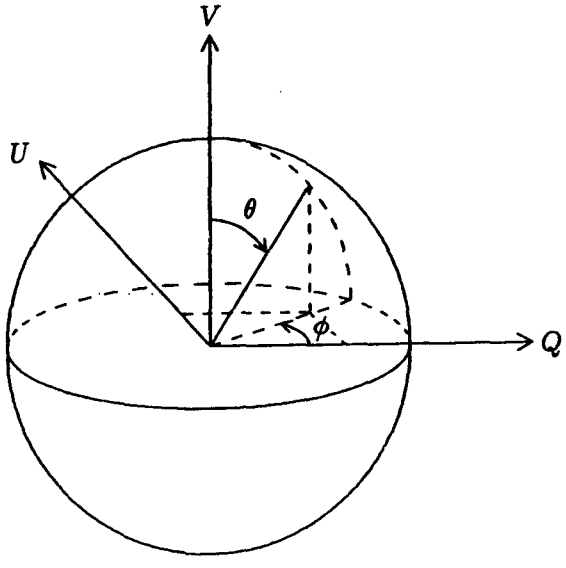


Figure 1.4.4 Poincaré Sphere.

Again we see from (14) that $I^2 = Q^2 + U^2 + V^2$.

Equation (14) suggests a simple geometrical representation of all states of polarization by recognizing that Q, U, and V can be regarded as the rectangular components of a point on a sphere with radius I [Fig. 1.4.4]. We define, in the spherical coordinate system, $\theta = \pi/2 - 2\beta$ and $\phi = 2\alpha$. As seen from (5), positive β is for right-hand polarization which is represented by points on the upper hemisphere. On the lower hemisphere, the points correspond to left-hand polarization. The north pole represents right-hand circular polarization and the south pole represents left-hand circular polarization. The sphere is called the *Poincaré sphere*. Figure 1.4.3 is seen to be a planar projection of the Poincaré sphere with the plane and the sphere touching each other at $Q = I$. The equator is mapped into the horizontal axis in Figure 1.4.3.

b. Partial Polarization

Radiation from many natural and man-made sources consists of

field components that fluctuate with time. We write

$$E_h = e_h(t)\cos\Big(\omega t - \psi_h(t)\Big)$$
$$E_v = e_v(t)\cos\Big(\omega t - \psi_v(t)\Big)$$

The wave is quasi-monochromatic when $e_h(t)$, $e_v(t)$, $\psi_h(t)$, and $\psi_v(t)$ are slowly-varying compared with $\cos\omega t$. The Stokes parameters are now defined by a time-average procedure denoted with the brackets $< >$.

$$< E_h^2(t) >= \frac{1}{T}\int_0^T dt\,[E_h(t)]^2$$

The Stokes parameters are

$$I = I_h + I_v = \frac{1}{\eta}\left(< E_h^2 > + < E_v^2 >\right)$$
$$Q = I_h - I_v = \frac{1}{\eta}\left(< E_h^2 > - < E_v^2 >\right) = I < \cos 2\alpha \cos 2\beta >$$
$$U = \frac{2}{\eta} < E_h E_v \cos\psi > = I < \sin 2\alpha \cos 2\beta >$$
$$V = \frac{2}{\eta} < E_h E_v \sin\psi > = I < \sin 2\beta >$$

For completely unpolarized waves, E_h and E_v are uncorrelated and we have $I =$ total Poynting power and $Q = U = V = 0$. For completely polarized waves we have $I^2 = Q^2 + U^2 + V^2$. For partially polarized waves it can be shown that $I^2 \geq Q^2 + U^2 + V^2$. With the Poincaré sphere of radius I, the partially polarized waves correspond to points inside the sphere.

In concluding this section on wave polarization, two remarks are in order. First, in defining the handedness of polarizations, we have taken the viewpoint of staying at a fixed point in space and observing the rotation of the tip of the electric field vector $\overline{E}$. If we take the other viewpoint of examining the spatial variation of $\overline{E}$ at fixed times, we shall find that a left-handed circularly polarized wave gives size to a right-handed helix in space [Problem 1.12]. Secondly, the polarization is defined according to the space and time variations of the $\overline{E}$ vector. As we shall see in Chapter II, it is imperative that we define polarization in terms of $\overline{D}$ when anisotropic and bianisotropic media are involved.

This is because in isotropic media $\overline{E}$ is perpendicular to $\overline{k}$, $\overline{k}\cdot\overline{E}=0$, while in non-isotropic media it is $\overline{k}\cdot\overline{D}=0$. This also suggests that wave polarization can be defined in terms of the field vector $\overline{B}$.

1.5 Boundary Conditions

The Maxwell equations have been written in differential form. They must be supplemented with boundary conditions and initial conditions wherever derivatives do not exist. The boundary conditions can be derived from the integral form of Maxwell's equations. First we consider the integration of a vector field $\overline{A}$ over a volume V enclosed by a surface S with surface normals $\hat{s}$. The following formulas are useful:

$$\iiint dV\,\nabla\cdot\overline{A}=\oiint dS\,\hat{s}\cdot\overline{A} \tag{1a}$$

$$\iiint dV\,\nabla\times\overline{A}=\oiint dS\,\hat{s}\times\overline{A} \tag{1b}$$

where (1a) is the familiar Gauss theorem which relates integration of the divergence of the vector field $\overline{A}$ over the volume V to the integration of the field over the surface S enclosing V. Equation (1b) is derived from (1a) by noting that $\nabla\cdot(\overline{C}\times\overline{A})=-\overline{C}\cdot\nabla\times\overline{A}$ where $\overline{C}$ is a constant vector independent of position. Applying the Gauss theorem (1a) to $\nabla\cdot(\overline{C}\times\overline{A})$, we obtain (1$b$) dot-multiplied by $\overline{C}$ on both sides. Letting $\overline{C}$ be an arbitrary vector, the result is then (1b).

Now consider an interface separating regions 1 and 2 [Fig. 1.5.1]. Assume a small pillbox volume across the interface. Integrating (1.1.1)–(1.1.4) over the volume and applying (1), we obtain

$$\oiint dS\,\hat{s}\times\overline{E}+\iiint dV\frac{\partial}{\partial t}\overline{B}=0 \tag{2}$$

$$\oiint dS\,\hat{s}\times\overline{H}-\iiint dV\frac{\partial}{\partial t}\overline{D}=\iiint dV\,\overline{J} \tag{3}$$

$$\oiint dS\,\hat{s}\cdot\overline{B}=0 \tag{4}$$

$$\oiint dS\,\hat{s}\cdot\overline{D}=\iiint dV\,\rho \tag{5}$$

These are Maxwell's equations in integral form.

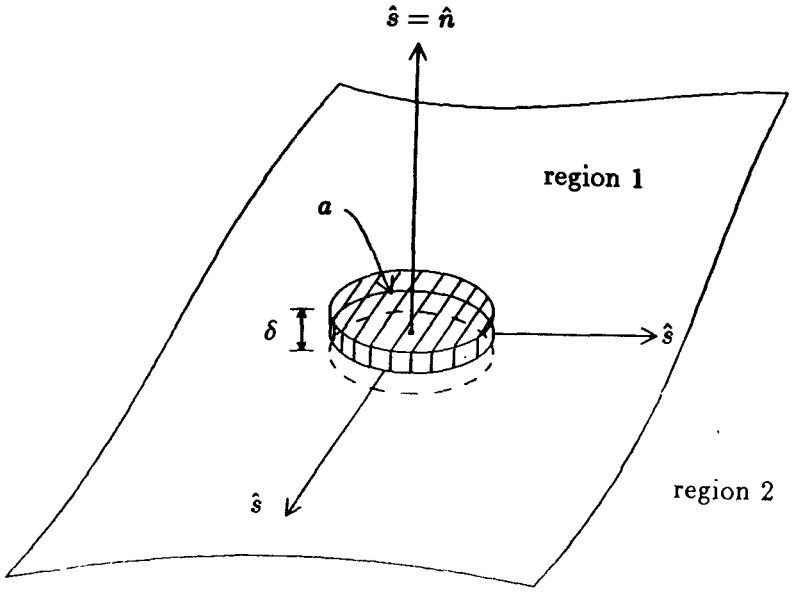


Figure 1.5.1 Pillbox for derivation of boundary conditions.

a. Stationary Boundaries

If we assume that the boundary surface is not in motion, then for the terms involving partial derivatives with time, $\partial/\partial t$ can be moved to the outside of the integral. Since the integration is over the volume, the result is a function of time only, and the partial derivatives become total derivatives. Therefore, for stationary boundary surfaces, Maxwell's equations in integral form become

$$\oiint dS\, \hat{s} \times \overline{E} = -\frac{d}{dt} \iiint dV\, \overline{B} \tag{6}$$

$$\oiint dS\, \hat{s} \times \overline{H} = \frac{d}{dt} \iiint dV\, \overline{D} + \iiint dV\, \overline{J} \tag{7}$$

$$\oiint dS\, \hat{s} \cdot \overline{B} = 0 \tag{8}$$

$$\oiint dS\, \hat{s} \cdot \overline{D} = \iiint dV\, \rho \tag{9}$$

Now we let the volume of the pillbox approach zero in such a manner that the thickness of the ribbon side, δ, goes to zero before the top

and bottom areas a shrink to a point. We dispose of terms of the order of δ.

The field vectors $\overline{E}$, $\overline{B}$, $\overline{D}$, and $\overline{H}$ are assumed to be finite but may be discontinuous across the boundary. The volume current and charge densities $\overline{J}$ and $\overline{\rho}$, however, may be infinite, such as on the surface of a perfect conductor, where we can define the surface current density $\overline{J}_s = \delta\overline{J}$ in the limit as $\delta \to 0$ and $\overline{J} \to \infty$, and define the surface charge density $\rho_s = \delta\rho$ in the limit as $\delta \to 0$ and $\rho \to \infty$. The surface current density has dimension amp/m and the surface charge density has dimension coulomb/m^2.

First we see that the terms involving time derivatives in (6) and (7) drop out because they are proportional to δ. We then consider the right-hand sides of (7) and (9) which become $\delta a\overline{J}$ and $\delta a\rho$, respectively. If $\overline{J}$ and ρ are finite, both terms will be zero because they are proportional to δ. When there are surface charges and currents at the boundary, the right-hand sides of (7) and (9) become $a\overline{J}_s$ and $a\rho_s$. We then see that the surface integral terms involving cross and dot products by $\hat{s}$ will be dropped except when $\hat{s}$ is in the directions $\hat{n}$ or $-\hat{n}$. After cancelling a on both sides of the equations, we obtain from (6)–(9) the following boundary conditions

$$\hat{n} \times (\overline{E}_1 - \overline{E}_2) = 0 \tag{10}$$

$$\hat{n} \times (\overline{H}_1 - \overline{H}_2) = \overline{J}_s \tag{11}$$

$$\hat{n} \cdot (\overline{B}_1 - \overline{B}_2) = 0 \tag{12}$$

$$\hat{n} \cdot (\overline{D}_1 - \overline{D}_2) = \rho_s \tag{13}$$

where subscripts 1 and 2 denote fields in regions 1 and 2, respectively. Essentially the boundary conditions state that the tangential components of $\overline{E}$ and the normal components of $\overline{B}$ are continuous across the boundary; the discontinuity of the tangential components of $\overline{H}$ is equal to the surface current density $\overline{J}_s$; and the discontinuity of the normal components of $\overline{D}$ is equal to the surface charge density ρ_s.

b. Moving Boundaries

The boundary conditions as stated in (10)–(13) are for stationary boundaries. When the boundary surface is moving, the partial time derivatives can no longer be commuted with the volume integral in (2)–(3). To derive the boundary conditions at moving boundaries, we let the pillbox move with the boundary surface. In accordance with kinematic

theory, for a moving volume with velocity $\overline{v}$ [Problem P1.15],

$$\frac{d}{dt}\iiint dV\,\overline{A} = \iiint dV\,\frac{\partial}{\partial t}\overline{A} + \oiint dS\,(\hat{s}\cdot\overline{v})\,\overline{A} \tag{14}$$

where $\overline{A}$ denotes any vector field. The surface integration term accounts for the motion of the boundary. We see that for moving boundaries, Maxwell's equations in integral form (2)–(5) become

$$\oiint dS\,[\hat{s}\times\overline{E} - (\hat{s}\cdot\overline{v})\overline{B}] = -\frac{d}{dt}\iiint_V dV\,\overline{B} \tag{15}$$

$$\oiint dS\,[\hat{s}\times\overline{H} + (\hat{s}\cdot\overline{v})\overline{D}] = \frac{d}{dt}\iiint_V dV\,\overline{D} + \iiint_V dV\,\overline{J} \tag{16}$$

$$\oiint dS\,(\hat{s}\cdot\overline{B}) = 0 \tag{17}$$

$$\oiint dS\,(\hat{s}\cdot\overline{D}) = \iiint dV\,\rho \tag{18}$$

We shrink the pillbox in the same manner as before such that terms of the order of δ are disposed of. The boundary conditions now become

$$\hat{n}\times(\overline{E}_1 - \overline{E}_2) - (\hat{n}\cdot\overline{v})(\overline{B}_1 - \overline{B}_2) = 0 \tag{19}$$

$$\hat{n}\times(\overline{H}_1 - \overline{H}_2) + (\hat{n}\cdot\overline{v})(\overline{D}_1 - \overline{D}_2) = \overline{J}_s \tag{20}$$

$$\hat{n}\cdot(\overline{B}_1 - \overline{B}_2) = 0 \tag{21}$$

$$\hat{n}\cdot(\overline{D}_1 - \overline{D}_2) = \rho_s \tag{22}$$

They reduce to (10)–(13) when $\overline{v} = 0$. It is important to note that they also reduce to (10)–(13) when $\hat{n}\cdot\hat{v} = 0$. Therefore, when the velocity is parallel to the interface, the boundary conditions of such a moving boundary are identical to those of a stationary boundary.

PROBLEMS

Problem P1.1

The constitutive relation $\overline{D} = \overline{\overline{\epsilon}}\cdot\overline{E}$ can also be expressed in terms of a "free-space" part $\epsilon_o\overline{E}$ and a polarization vector $\overline{P}$ characterizing the properties of the material. We write

$$\overline{D} = \epsilon_o\overline{E} + \overline{P}$$

In the case of induced dipole moments, the polarization $\overline{P}$ is proportional to the polarizability per unit volume $N\alpha$, where N is the number of dipoles per unit volume

$$\overline{P} = N\alpha\overline{E}_{loc}$$

The local electric field $\overline{E}_{loc}$ at the place of the induced dipole comprises the field $\overline{E}$ and the field caused by the surrounding dipoles $\overline{E}_{loc} = \overline{E} + \overline{E}_{SD}$. If a spherical cavity of radius a is assumed to surround the dipole, we write

$$\overline{E}_{SD} = \begin{cases} -E_0 \left(\frac{r}{a}\right)^3 (\hat{r}\, 2\cos\theta + \hat{\theta}\sin\theta) & r \geq a \\ E_0(\hat{r}\cos\theta - \hat{\theta}\sin\theta) & r \leq a \end{cases}$$

Show that the tangential component of $\overline{E}_{SD}$ is continuous at $r = a$ and the discontinuity of the normal component of $\epsilon_o\overline{E}_{SD}$ at $r = a$ gives rise to the polarization surface charge. Thus deduce that

$$\overline{E}_{loc} = \overline{E} + \frac{\overline{P}}{3\epsilon_o}$$

Show that

$$\frac{\epsilon}{\epsilon_o} = \frac{1 + (2N\alpha/3\epsilon_o)}{1 - (N\alpha/3\epsilon_o)}$$

This is the well-known Clausius-Mossotti (or Lorentz-Lorenz) formula.

Problem P1.2

Similar to the expression of the constitutive relation $\overline{D} = \overline{\overline{\epsilon}} \cdot \overline{E} = \epsilon_o\overline{E} + \overline{P}$, the constitutive relation $\overline{B} = \overline{\overline{\mu}} \cdot \overline{H}$ can also be represented in terms of a "free-space" part $\mu_o\overline{H}$ and a magnetization vector $\overline{M}$ such that

$$\overline{B} = \mu_o\overline{H} + \mu_o\overline{M}$$

Notice that while $\overline{P}$ has the same dimension as $\overline{D}$, $\overline{M}$ has the same dimension as $\overline{H}$.

In the case of media possessing permanent moments, the polarization $\overline{P}$ and the magnetization $\overline{M}$ are given classically by the Langevin equation

$$L(x) = \coth x - \frac{1}{x}$$

For a paramagnetic material with magnetic moments Nm,

$$M = NmL\left(\frac{mH}{kT}\right)$$

where $k = 1.38 \times 10^{-23}$ Joule/Kelvin is Boltzmann's constant, and T is the absolute temperature in Kelvins. Show that in the low-field limit, since $mH \ll kT$, the medium is linear.

Problem P1.3

For each of the following constitutive relations, state whether the given medium is

(1) Isotropic/anisotropic/bianisotropic,
(2) Linear/nonlinear,
(3) Spatially/temporally dispersive,
(4) Homogeneous/inhomogeneous.

(a) Cholesteric liquid crystals can be modelled by a spiral structure with constitutive relations given by

$$\overline{D} = \begin{pmatrix} \epsilon(1+\delta\cos Kz) & \epsilon\delta\sin Kz & 0 \\ \epsilon\delta\sin Kz & \epsilon(1-\delta\cos Kz) & 0 \\ 0 & 0 & \epsilon_z \end{pmatrix} \cdot \overline{E}$$

where the spiral direction is along the z axis.

(b) In view of the optical activities in quartz crystals, the constitutive relation for a quartz crystal is phenomenologically described as

$$E_j = \kappa_{ij} D_i + \frac{1}{\mu_o \epsilon_o} G_{ij} \frac{\partial}{\partial t} B_i$$

$$H_j = \frac{1}{\mu_o} B_j - \frac{1}{\mu_o \epsilon_o} G_{ij} \frac{\partial}{\partial t} D_i$$

(c) When a magnetic field $\overline{B}_0$ is applied to a conductor carrying a current, an electric field $\overline{E}$ is developed. This is called the *Hall effect*. Assuming the conduction carrier drifts with a mean velocity $\overline{v}$ proportional to $R\sigma\overline{E}$, the constitutive relation that takes care of the Hall effect is given by

$$\overline{J} = \sigma\left(\overline{E} + R\sigma\overline{E} \times \overline{B}_0\right)$$

where σ is the conductivity and R is the Hall coefficient. For copper, $\sigma \approx 6.7 \times 10^7$ mho/m and $R \approx -5.5 \times 10^{-11}$ m^3/coul.

(d) The phenomenon of natural optical activity can be explained with the use of the constitutive relation [Landau and Lifshitz, 1957]

$$D_i = \epsilon_{ij} E_j + \gamma_{ijk} \frac{\partial E_j}{\partial x_k}$$

where ϵ_{ij} and γ_{ijk} are functions of frequency and $\gamma_{ijk} = -\gamma_{jik}$.

(e) The phenomenon of pyroelectricity in a crystal is observed when it is heated. The constitutive relation for a *pyroelectric material* can be written as

$$\overline{D} = \overline{D}_0 + \bar{\bar{\epsilon}} \cdot \overline{E}$$

where a spontaneous term $\overline{D}_0$ exists even in the absence of an external field.

(f) The phenomenon in which dipole moments are induced in a crystal by mechanical stress is called *piezoelectricity*. A piezoelectric material is characterized by a piezoelectric tensor $\gamma_{i,kl} = \gamma_{i,lk}$ such that

$$D_i = D_{0i} + \epsilon_{ik} E_k + \gamma_{i,kl} s_{kl}$$

where s_{kl} is the stress tensor to second order in electric fields. All pyroelectric media are also piezoelectric.

(g) An isotropic dielectric can exhibit the *Kerr effect* when placed in an electric field. In this case the permittivity can be written as

$$\epsilon_{ij} = \epsilon \delta_{ij} + \sigma E_i E_j$$

where ϵ is the unperturbed permittivity. The principal axis of ϵ_{ij} coincides with the electric field.

(h) In an electrooptical material that exhibits *Pockel's effect* the constitutive relation can be written as

$$D_i = \epsilon_{ij} E_j + \sigma_{ijk} E_j E_k$$

where $\sigma_{ijk} = \sigma_{jik}$ is a third-rank tensor symmetrical in i and j, and therefore has 18 independent elements.

Problem P1.4

Electromagnetic waves satisfy all of Maxwell's equations. Consider, in free space, the following electric field vectors:

$$\overline{E}_1 = \hat{x}\cos(\omega t - kz)$$
$$\overline{E}_2 = \hat{z}\cos(\omega t - kz)$$
$$\overline{E}_3 = (\hat{x} + \hat{z})\cos\left(\omega t + k|x - z|/\sqrt{2}\right)$$
$$\overline{E}_4 = (\hat{x} + \hat{z})\cos\left(\omega t + k|x + z|/\sqrt{2}\right)$$
$$\overline{E}_5 = (\hat{x} + \hat{z})\cos(\omega t + ky)$$

(a) Do these electric field vectors satisfy the wave equation

$$\left(\nabla^2 - \mu_o\epsilon_o\frac{\partial^2}{\partial t^2}\right)\overline{E} = 0$$

What is the relationship between ω and k?

(b) Find the corresponding magnetic field vectors for each of the five electric field vectors.

(c) Which of the five fields qualify as electromagnetic waves? For those not qualified as electromagnetic waves, you should state which of Maxwell's equations are violated. Show that for those qualified as electromagnetic waves, the electric and magnetic field vectors are perpendicular to the direction of propagation.

Problem P1.5

Derive an equation for $\overline{E}$ for a conducting medium. Show that

$$\overline{E}(z,t) = \hat{x}e^{-k_I z}\cos(k_R z - \omega t)$$

is a solution to the equation by finding the values for k_I and k_R.

Problem P1.6

The Earth receives over all frequency band about 1.5 kW /m^2 of power from the sun.

(a) The Earth-Sun distance is 150 Gm, how long does it take the sunlight to reach the Earth ?

(b) The Earth radius is 6400 km, what is the total power received by the Earth ?

(c) Assume the Sun's mass is 2×10^{30} kg which converts to radiation according to mc^2 at 1 percent efficiency, how long can the Sun radiate at the present level ?

(d) The Sun radiates $10^{-20}\,\mathrm{w\,m^{-2}\,Hz^{-1}}$ at 3 GHz. Assuming constant power level over 1 GHz bandwidth, what is the Poynting power density and the corresponding electric field amplitude ?

Problem P1.7

The *solar wind* is a high-conductivity plasma which is emitted radially from the surface of the Sun. Let us calculate the flux of electromagnetic energy in the solar wind at the orbit of the Earth.

In the plane of the Earth's orbit, the magnetic field of the Sun is approximately radial, pointing outwards in certain regions and inwards in others. This field is "frozen" in the high-conductivity plasma. Since the Sun rotates (with a period of 27 days), and the plasma has a radial velocity, the lines of $\overline{B}$ are in fact Archimedes spirals ($r = a\theta$ in polar coordinates) and, at the Earth, they form an angle of about 45° with the Sun-Earth direction. This is the so-called *garden hose* effect.

At the orbit of the Earth the solar wind has a density of about 10^7 proton-masses/m^3 and a velocity of about 4×10^5 m/sec, while the magnetic field of the Sun is about 5×10^{-9} tesla (webers/m^2).

(a) First show that, in an electrically neutral ($\rho = 0$) and nonmagnetic fluid of conductivity σ and velocity $\overline{v}$, Maxwell's equations become

$$\nabla \cdot \overline{D} = 0$$

$$\nabla \cdot \overline{B} = 0$$

$$\nabla \times \overline{E} = -\frac{\partial \overline{B}}{\partial t}$$

$$\nabla \times \overline{B} = \mu_o \left\{ \sigma(\overline{E} + \overline{v} \times \overline{B}) + \epsilon_o \frac{\partial \overline{E}}{\partial t} \right\}$$

the polarization currents being negligibly small compared to the conduction currents.

Note that, for an infinite conductivity,

$$\overline{E} = -\overline{v} \times \overline{B}$$

This is a satisfactory approximation for the solar wind.

(b) Show that the component of $\overline{v}$ which is normal to $\overline{B}$ is

$$\overline{v}_n = \frac{1}{B^2}\overline{B} \times (\overline{v} \times \overline{B})$$

(c) Show that the Poynting vector of the solar wind is

$$\overline{S} = \frac{B^2}{\mu_o}\overline{v}_n$$

Numerically it is approximately equal to 4×10^{-9} times the average value of the Poynting vector of the solar radiation, which is about 1.4kW/m^2. The Poynting vector of the solar wind is normal to the local $\overline{B}$ and it points at an angle of $45°$ away from the Sun-Earth direction.

(d) Compare the relative magnitudes of the kinetic, electric, and magnetic energy densities. Which is largest?

Problem P1.8

The interpretation of $\overline{S} = \overline{E} \times \overline{H}$ as power flow per unit area at a point in space is a very useful and experimentally verified concept in electromagnetic wave theory. However, strict adherence to such interpretation in total disregard of (1.3.1) may lead to paradoxical results:

(a) Consider a charged particle placed next to a permanent magnet. Is there power flow due to the electric field $\overline{E}$ generated by the charged particle and the magnetic field $\overline{H}$ generated by the magnet? Show that the net power flow is zero for static electric and magnetic fields in the absence of currents.

(b) Consider a current density J flowing through a resistor of conductivity σ, length l, and circular cross-section with radius a. The electric field is $E = J/\sigma$, and the magnetic field $H = aJ/2$. Show that the total Poynting power flow into the resistor is $\pi a^2 l J^2/\sigma$. Draw similarity to the potential energy of a raised weight, state your view as to the location of the source of energy giving rise to the Poynting power flow.

Problem P1.9

Consider an electromagnetic wave propagating in the $\hat{z}$ direction and described by

$$E = \hat{x}E_x \cos(kz - \omega t + \psi_x) + \hat{y}E_y \cos(kz - \omega t + \psi_y)$$

where E_x, E_y, ψ_x, and ψ_y are all real numbers.

(a) Let $E_x = 2$, $E_y = 1$, $\psi_x = \pi/2$, $\psi_y = \pi/4$. Draw a picture and show what the polarization is.

(b) Let $E_x = 1$, $E_y = \psi_x = 0$. This is a linearly polarized wave. Prove that it can be expressed as the superposition of a right-hand circularly polarized wave and a left-hand circularly polarized wave.

(c) Let $E_x = 1$, $\psi_x = \pi/4$, $\psi_y = -\pi/4$, $E_y = 1$. This is a circularly polarized wave. Prove that it can be decomposed into two linearly polarized waves.

Problem P1.10

Sun navigation was first observed in 1911. It was found that some species of ants, horseshoe crabs, honeybees, etc., are sensitive to polarized light. These creatures can navigate as long as there is a small patch of blue sky. The sky polarization depends upon the angle ϕ between the sun's rays to a particular point in the sky and an observer's line of sight to the same point. This is because sunlight is scattered by air molecules, which behave like small dipole antennas when irradiated by sunlight. The scattered electric field $\overline{E}_s$ of the dipole is seen to be polarized in planes containing the induced dipole. The magnitude of the scattered field is maximum in the direction perpendicular to the dipole axes. (This field of a dipole antenna will be discussed quantitatively in Chapter IV.)

Draw a picture to show that when the sun is viewed directly the light is unpolarized and that it becomes more linearly polarized when looking at other parts of the sky. Show that as ϕ approaches 90° the polarization will tend to be linear [*Scientific American*, July 1955].

Problem P1.11

Assume $E_v(t)$, $E_h(t)$ and $\psi(t)$ remain constant for a fractional time interval t_n and let T be subdivided into $t_1, t_2, \ldots t_N$. The time-average can be written in a summation form such that during the

fractional time intervals $t_1, t_2, t_3, \ldots$ adding up to T,

$$I_h = \frac{1}{\eta T} \sum_n t_n |E_{hn}|^2$$

$$I_v = \frac{1}{\eta T} \sum_n t_n |E_{vn}|^2$$

$$U = \frac{2}{\eta T} \sum_n t_n |E_{vn}|^2 A_n \cos \psi_n$$

$$V = \frac{2}{\eta T} \sum_n t_n |E_{vn}|^2 A_n \sin \psi_n$$

where

$$A_n = \frac{|E_{hn}|}{|E_{vn}|}$$

denotes the ratio of $|E_h|$ and $|E_v|$ in the time interval t_n.

(a) Show that

$$4 I_v I_h = \frac{4}{\eta^2 T^2} \left\{ \sum_n t_n^2 A_n^2 |E_{vn}|^4 + \sum_{(n,m)} t_n t_m |E_{vn}|^2 |E_{vm}|^2 \right.$$
$$\left. (A_m^2 + A_n^2) \right\}$$

$$U^2 + V^2 = \frac{4}{\eta^2 T^2} \left\{ \sum_n t_n^2 A_n^2 |E_{vn}|^4 + 2 \sum_{(n,m)} t_n t_m |E_{vn}|^2 |E_{vm}|^2 \right.$$
$$\left. A_n A_m \cos(\psi_n - \psi_m) \right\}$$

where (n, m) indicates that the summation is to be extended over all distinct pairs of n and m. We find

$$4 I_v I_h - (U^2 + V^2) = \frac{4}{\eta^2 T^2} \sum_{(n,m)} t_n t_m |E_{vn}|^2 |E_{vm}|^2$$
$$\left\{ A_m^2 + A_n^2 - 2 A_n A_m \cos(\psi_n - \psi_m) \right\}$$

The right-hand side is always positive. Hence

$$4 I_v I_h \geq U^2 + V^2$$

or

$$I^2 = (I_h + I_v)^2 \geq Q^2 + U^2 + V^2$$

The equal sign holds when $\psi_n = \psi_m$ and $A_n = A_m$ for all n and m, which means that the amplitude ratio and phase difference of E_v and E_h stay constant. This is the case for elliptical polarization.

(b) For polarized waves

$$\begin{aligned} I &= I_h + I_v \\ Q &= I_h - I_v = I\cos 2\alpha \cos 2\beta \\ U &= I\sin 2\alpha \cos 2\beta \\ V &= I\sin 2\beta \end{aligned}$$

Show that when the wave is right-handed circularly polarized $Q = U = 0$ and $V = I$, when it is left-hand circularly polarized, $Q = U = 0$ and $V = -I$, and when the wave is linearly polarized, $V = 0$.

Problem P1.12

In 1.2, wave motion is viewed by either taking a series of still pictures at several fixed times or by making observations at a fixed point in space. The definition of polarization in 1.3 is discussed from the second point of view. Let us look at polarization from the first point of view in this problem.

Consider the following wave solution

$$\overline{E}(\overline{r}, t) = E_0[\hat{x}\cos(kz - \omega t) + \hat{y}\sin(kz - \omega t)]$$

Show that it satisfies the wave equation and that it represents a left-hand circularly polarized wave according to the definition for polarization in 1.3. Show that at time $t = 0$, the locus of the tip of the electric field vectors along the z axis forms a ***right-handed helix*** [Fig. P1.12].

The parametric equation of a helix is

$$\begin{aligned} x &= R\cos\left(\frac{2\pi}{p}z\right) \\ y &= R\sin\left(\frac{2\pi}{p}z\right) \end{aligned}$$

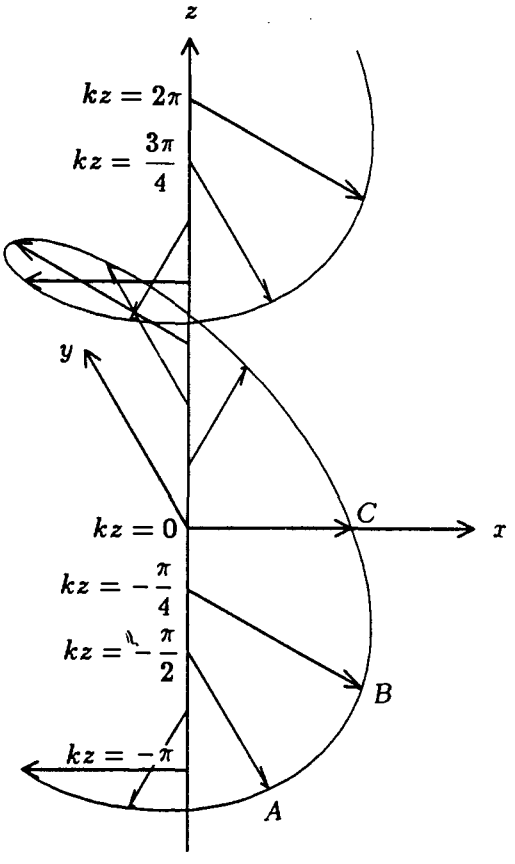


Figure P1.12

where p is the pitch of the helix. Find the pitch and show that the helix advances along the $\hat{z}$ direction without rotating.

Problem P1.13

Consider an electric field which is linearly polarized along the x axis,

$$E_i = \hat{x} E_0 \cos(\omega t - kz)$$

where $E_0 = 1$ volt/m.

A piece of polaroid with optic axis along $\hat{x}$ is placed perpendicular to the z axis at $z = a$, and another piece of polaroid with optic axis along $\hat{y}$ is placed perpendicular to the z axis at $z = c$. The

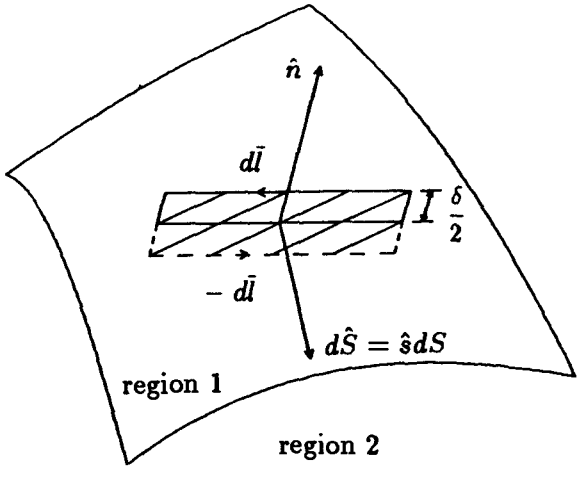


Figure P1.14

transmitted fields ($z > c$) will be zero. Now a third polaroid is inserted at $z = b$ with $a < b < c$, with the angle θ between the optic axis and the x axis. When θ is varied, we can have some transmission of energy at $z > c$.

Calculate and sketch the Poynting power density transmitted as a function of θ. (Assume 100% power transmission when the incident field is aligned with the optic axis of a polaroid, and neglect reflections at the surface of the polaroids.) What is the angle θ such that the transmitted power is maximum?

Problem P1.14

The boundary conditions for the two curl equations can be derived, instead of using a small pillbox as shown in the text, by considering a ribbon-like surface as shown in Figure P1.14. Integrating over the surface of the ribbon area, Faraday's law and Ampere's law become

$$\oint d\bar{l} \cdot \overline{E} = -\frac{d}{dt} \iint dS\, \hat{s} \cdot \overline{B}$$

$$\oint d\bar{l} \cdot \overline{H} = \frac{d}{dt} \iint dS\, \hat{s} \cdot \overline{D} + \iint dS\, \hat{s} \cdot \overline{J}$$

Show that (1.5.10)–(1.5.11) follow as we let the ribbon area approach zero in such a manner that δ goes to zero first and the terms involving δ are discarded.

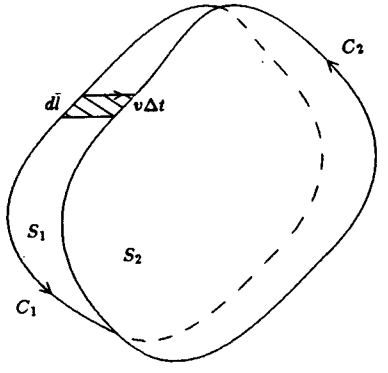


Figure P1.15

Problem P1.15

For moving boundary surfaces, the boundary conditions for the two curl equations can be derived by using the same ribbon-like area shown in the preceding problem but now moving with the surface. First we derive a kinematic formula for the time derivative of a moving surface.

Consider at time t a surface S_1 bounded by the contour C_1 [Fig. P1.15]. Let $\bar{v}$ be the instantaneous velocity of the element $d\bar{S}$ of the surface. The surface S_1 together with the contour C_1 may change shape with time, as $\bar{v}$ need not be a constant for all elements of S_1. At time $t + \Delta t$, S_1 and C_1 become S_2 and C_2 [Fig. P1.15].

(a) Applying the divergence theorem to the volume bounded by $\overline{S}_1$, $\overline{S}_2$, and the ribbon area $d\bar{l} \times \bar{v}\Delta t$ [Fig. P1.15]

$$\iint_S (d\bar{S} \cdot \bar{v}\Delta t)(\nabla \cdot \bar{A}) = \iint_{S_2} d\bar{S}_2 \cdot \bar{A} - \iint_{S_1} d\bar{S}_1 \cdot \bar{A} + \oint_C (d\bar{l} \times \bar{v}\Delta t) \cdot \bar{A}$$

show that the total time derivative of a vector field $\bar{A}$ integrated over the surface is

$$\frac{d}{dt} \iint_S d\bar{S} \cdot \bar{A} = \iint_S d\bar{S} \cdot \left[\frac{\partial \bar{A}}{\partial t} + \nabla \times (\bar{A} \times \bar{v}) + \bar{v} \nabla \cdot \bar{A} \right]$$

The rate at which the flux of $\bar{A}$ through S changes is seen to depend on three processes. The first term is due to the time rate of change of $\bar{A}$ for a stationary contour. The second term accounts for the contribution due to the flux crossing the surface generated

Figure P1.16

by the motion of the contour C. The last term arises when the surface moves through the regions of the sources for the flux $\overline{A}$. When $\overline{A}$ is identified as the magnetic induction $\overline{B}$, the last term will be zero.

(b) Casting Maxwell's two curl equations in integral form

$$\oint d\bar{l} \cdot (\overline{E} + \overline{v} \times \overline{B}) = -\frac{d}{dt} \iint dS\, \hat{s} \cdot \overline{B}$$

$$\oint d\bar{l} \cdot (\overline{H} - \overline{v} \times \overline{D}) = \frac{d}{dt} \iint dS\, \hat{s} \cdot \overline{D} + \iint dS\, \hat{s} \cdot (\overline{J} - \overline{v}\rho)$$

and appling the integrals to a ribbon-like area moving with the boundary surface, derive the boundary conditions in (1.5.19–22).

Problem P1.16

In the derivation of the jump conditions for moving boundaries, show that

$$\frac{d}{dt} \iiint dV\, \overline{A} = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left\{ \iiint_{t+\Delta t} dV\, \overline{A}(t + \Delta t) - \iiint_t dV\, \overline{A}(t) \right\}$$

$$= \iiint dV \frac{\partial}{\partial t} \overline{A} + \iint dS\, (\hat{s} \cdot \overline{v}) \overline{A}$$

as given in (1.5.14). In the text, the jump boundary conditions are obtained by assuming that the pillbox moves with the boundary.

If the pillbox is stationary while the boundary is moving, then $v = 0$ in (1.5.15) as the velocity v refers to the pillbox. We should be able to derive the boundary condition (1.5.19) from (1.5.6). Consider (1.5.6) where

$$\oiint dS\, \hat{s} \times \overline{E} = -\frac{d}{dt} \iiint dV\, \overline{B}$$

Refer to Figure P1.16. The boundary surface has been moved a distance $v\Delta t$ from time t to time $t + \Delta t$. Show that

$$\iiint_t dV\, \overline{B} = \overline{B}_1 a(\delta - \delta_0) + \overline{B}_2 a \delta_0$$

and that

$$\iiint_{t+\Delta t} dV\, \overline{B} - \iiint_{t} dV\, \overline{B} = \overline{B}_2\, av\Delta t - \overline{B}_1 av\Delta t$$

where a is the area of the top or bottom side of the pillbox and $\overline{B}_1$ and $\overline{B}_2$ denote the $\overline{B}$ field in medium 1 and medium 2, respectively. From (1.5.6) deduce that

$$\hat{n} \times (\overline{E}_1 - \overline{E}_2) - (\hat{n} \cdot \overline{v})(\overline{B}_1 - \overline{B}_2) = 0$$

which is (1.5.19). Following a similar procedure, derive (1.5.20)–(1.5.22).

II

PROPAGATION IN HOMOGENEOUS MEDIA

2.1 Time-Harmonic Fields
- a. Maxwell's Equations for Time-Harmonic Fields
- b. Constitutive Relations and Dispersive Media
- c. Gyrotropic Media
- d. Complex Poynting's Theorem
- e. Time-Average Poynting Power Vector
- f. Symmetry Conditions for Lossless Media

2.2 Plane Wave Solutions
- a. Phase and Group Velocities
- b. Evanescent Waves in Lossless Media
- c. Penetration Depth in Lossy Media

2.3 Plane Waves in Homogeneous Media and the kDB System
- a. kDB System
- b. Maxwell's Equations in kDB System

2.4 Plane Waves in Uniaxial Media
- a. Ordinary and Extraordinary Waves
- b. k Surfaces

2.5 Plane Waves in Gyrotropic Media
- a. Dispersion Relations and Characteristic Waves
- b. Faraday Rotation

2.6 Plane Waves in Bianisotropic Media
- a. Dispersion Relations and Characteristic Waves
- b. Optical Activity

2.7 Plane Waves in Nonlinear Media
- a. Second-Harmonic Generation (SHG)
- b. Phase Conjugation

Problems

2.1 Time-Harmonic Fields

For electromagnetic waves of a particular frequency in the steady state, the fields are time-harmonic and are known as monochromatic waves or continuous waves (CW). The CW cases are important for three reasons: (i) the CW assumption can be used to eliminate the time dependence in the Maxwell equations and thus considerably simplify the mathematics; (ii) once the CW case is solved and a sound understanding is developed for the frequency-domain phenomena, Fourier theory can be applied to study the time-domain phenomena; (iii) CW representation covers the whole spectrum of electromagnetic waves. Clearly, a thorough understanding of CW or the time-harmonic case is essential in the study of all electromagnetic wave phenomena.

For a time-harmonic field with angular frequency ω, we let

$$\overline{E}(\bar{r},t) = Re\{\overline{E}(\bar{r})\, e^{-i\omega t}\} \tag{1}$$

where Re denotes the real part of a complex quantity and $e^{-i\omega t}$ denotes the convention of time dependence used in this book. The electric field vector $\overline{E}(\bar{r})$ is now independent of time but it becomes a complex vector that is a function of position only. Notice that in this book we shall not use different symbols to distinguish real quantities such as $\overline{E}(\bar{r},t)$ in the time domain and complex quantities such as $\overline{E}(\bar{r})$ in the frequency domain. Their meanings should be clear from the context. In case of possible ambiguity, we shall explicitly indicate the complex field quantities to be functions of $\bar{r}$ only and the real time-domain fields to be functions of both $\bar{r}$ and t.

a. Maxwell's Equations for Time-Harmonic Fields

Similar definitions apply to other field quantities with $\overline{E}$ replaced by $\overline{H}$, $\overline{B}$, $\overline{D}$, $\overline{J}$, and ρ in (1). Substituting $\overline{E}(\bar{r},t)$ and $\overline{B}(\bar{r},t)$ in (1.1.1) we obtain

$$Re\left\{[\nabla \times \overline{E}(\bar{r}) - i\omega\overline{B}(\bar{r})]e^{-i\omega t}\right\} = 0 \tag{2}$$

This equation is true *for all time t*. When the real part of the complex quantity in the square brackets multiplied by all values of $e^{-i\omega t}$ is equal to zero, the complex quantity itself must be equal to zero. This is easily seen by considering $Re\{Ce^{-i\omega t}\} = 0$ where $C = C_R + iC_I$ denotes the complex quantity with both C_R and C_I real. Letting

$\omega t = 0$, we find $C_R = 0$. Letting $\omega t = \pi/2$, we find $C_I = 0$. It follows that $C = 0$. We therefore conclude that

$$\nabla \times \overline{E}(\bar{r}) - i\omega \overline{B}(\bar{r}) = 0 \tag{3}$$

Similar arguments apply to other Maxwell equations and we find that they become, if we omit writing the argument $\bar{r}$,

$$\nabla \times \overline{E} = i\omega \overline{B} \tag{4}$$

$$\nabla \times \overline{H} = -i\omega \overline{D} + \overline{J} \tag{5}$$

$$\nabla \cdot \overline{B} = 0 \tag{6}$$

$$\nabla \cdot \overline{D} = \rho \tag{7}$$

$$\nabla \cdot \overline{J} = i\omega\rho \tag{8}$$

The boundary conditions are identical to those given in (1.5.10)-(1.5.13) and (1.5.19)–(1.5.22) where the field quantities are now complex.

It is important to note that the complex field vectors are convenient representations of the real field vectors. To recover the real space–time-dependent field vector, we simply multiply the complex quantity by $e^{-i\omega t}$ and take its real part as shown in (1).

b. Constitutive Relations and Dispersive Media

Under the time harmonic representation, constitutive elements are, in general, complex numbers. Consider a conducting medium governed by Ohm's law $\overline{J}_c = \overline{\overline{\sigma}} \cdot \overline{E}$. From the Maxwell equation

$$\nabla \times \overline{H} = -i\omega \overline{D} + \overline{J}_c + \overline{J}_f \tag{9}$$

where $\overline{J}_f$ represents the source, we can absorb $\overline{J}_c$ in $\overline{D}$ by noting that for a general bianisotropic medium $\overline{D} = \overline{\overline{\epsilon}} \cdot \overline{E} + \overline{\overline{\xi}} \cdot \overline{H}$. We find

$$\nabla \times \overline{H} = -i\omega \left[\overline{\overline{\epsilon}} + \frac{i}{\omega}\overline{\overline{\sigma}}\right] \cdot \overline{E} - i\omega \overline{\overline{\xi}} \cdot \overline{H} + \overline{J}_f$$

Thus we define a new permittivity tensor

$$\overline{\overline{\epsilon}}_{new} = \overline{\overline{\epsilon}} + \frac{i}{\omega}\overline{\overline{\sigma}} \tag{10}$$

which is complex and accounts for the conductivity of the medium. When both $\bar{\bar{\epsilon}}$ and $\bar{\bar{\sigma}}$ are isotropic and real, for instance, the conductivity σ then constitutes the imaginary part of a complex permittivity $\epsilon + i\sigma/\omega$.

Time dispersion is a common phenomenon for most media in the presence of time-varying fields. As an example, the permittivity of water drops from $80\epsilon_o$ to approximately $1.8\epsilon_o$ as the frequency increases from static to the optical range. The reason for this decrease is that the alignment of water molecules, which possess permanent dipole moments, is much more ineffective at optical frequencies than in slowly varying fields. As another example, consider a plasma medium, which is a neutral ionized gas consisting of free electrons and positive ions. Since the ions are much heavier than the electrons, we assume that only the interaction between the free electrons and electromagnetic waves need be considered. Let the electron plasma comprising electrons with density N, electron mass 9.1×10^{-31} kg, and electron charge $q = -1.6 \times 10^{-19}$ coul. Under an applied electromagnetic wave field, an electron is subject to a force $\bar{f} = q(\bar{E} + \bar{v} \times \bar{B}) \approx q\bar{E}$ for $v/c \ll 1$. From Newton's second law $q\bar{E} = d(m\bar{v})/dt = -i\omega m\bar{v}$ under time-harmonic excitations. The volume current due to the electrons is $\bar{J}_p = Nq\bar{v}$, where N is the number of electrons/m^3. From the Maxwell equation

$$\nabla \times \bar{H} = -i\omega\bar{D} + \bar{J}_p + \bar{J}_f = -i\omega\epsilon(\omega)\bar{E} + \bar{J}_f$$

where

$$\epsilon(\omega) = \epsilon_o \left[1 - \frac{\omega_p^2}{\omega^2}\right] \tag{11}$$

and the plasma frequency ω_p is defined to be

$$\omega_p = \sqrt{\frac{Nq^2}{m\epsilon_o}} \approx 56.4\sqrt{N} \tag{12}$$

Equation (11) explicitly displays the frequency dependence of the permittivity ϵ. It is noted that ϵ is always less than ϵ_o.

c. *Gyrotropic Media*

An electron plasma as described in (11) becomes anisotropic when an external *dc* magnetic field $\bar{B}_0$ is applied. The permittivity tensor

becomes [Prob. P2.1]

$$\overline{\overline{\epsilon}} = \begin{pmatrix} \epsilon & i\epsilon_g & 0 \\ -i\epsilon_g & \epsilon & 0 \\ 0 & 0 & \epsilon_z \end{pmatrix} \tag{13}$$

if we assume that $\overline{B}_0$ is in the $\hat{z}$ direction. The magnetic field enters the constitutive parameters via the cyclotron frequency ω_c,

$$\omega_c = \frac{qB_0}{m}$$

The parameters in (13) are defined by

$$\epsilon = \epsilon_o \left[1 - \frac{\omega_p^2}{(\omega^2 - \omega_c^2)}\right]$$

$$\epsilon_g = \epsilon_o \left[\frac{-\omega_p^2 \omega_c}{\omega(\omega^2 - \omega_c^2)}\right]$$

$$\epsilon_z = \epsilon_o \left(1 - \frac{\omega_p^2}{\omega^2}\right)$$

In the case of an infinitely strong magnetic field, we have

$$\epsilon = \epsilon_o$$

$$\epsilon_g = 0$$

$$\epsilon_z = \epsilon_o \left(1 - \frac{\omega_p^2}{\omega^2}\right)$$

and the plasma becomes a ***uniaxial plasma.***

An anisotropic medium characterized by a hermitian permittivity tensor such as (13) is called ***gyroelectric*** or ***electrically gyrotropic.*** An anisotropic medium that is characterized by a hermitian permeability tensor $\overline{\overline{\mu}}$ such as

$$\overline{\overline{\mu}} = \begin{pmatrix} \mu & i\mu_g & 0 \\ -i\mu_g & \mu & 0 \\ 0 & 0 & \mu_z \end{pmatrix} \tag{14}$$

is called a ***gyromagnetic*** medium or a ***magnetically gyrotropic*** medium. An example is a ferrite subject to a dc magnetic field in the $\hat{z}$ direction

around which the magnetization of the ferrite precesses. Note that the gyrotropic elements in a gyrotropic medium, although imaginary, do not account for any loss.

d. Complex Poynting's Theorem

The complex Poynting's theorem is derived by dot-multiplying (2.1.4) by $\overline{H}^*$ and subtracting the complex conjugate of (2.1.5) dot-multiplied by $\overline{E}$. Making use of the identity $\overline{H}^* \cdot \nabla \times \overline{E} - \overline{E} \cdot \nabla \times \overline{H}^* = \nabla \cdot (\overline{E} \times \overline{H}^*)$, we obtain

$$\nabla \cdot (\overline{E} \times \overline{H}^*) = i\omega \left[\overline{B} \cdot \overline{H}^* - \overline{E} \cdot \overline{D}^* \right] - \overline{E} \cdot \overline{J}^* \tag{15}$$

The complex Poynting's vector $\overline{S}$ is defined to be

$$\overline{S} = \overline{E} \times \overline{H}^* \tag{16}$$

However, it is noted that, mathematically, $\overline{E} \times \overline{H}^*$ is not a uniquely defined quantity as far as Poynting's theorem is concerned. An arbitrary curl field $\nabla \times \overline{A}$ can be added to $\overline{E} \times \overline{H}^*$ without changing (15). Physically the complex vector $\overline{S}$ as defined in (16) has been identified as a complex power density vector.

The term $\overline{E} \cdot \overline{J}^* = \overline{E} \cdot (\overline{J}_c^* + \overline{J}_f^*)$ in (15) consists of two parts: one part due to the ohmic current $\overline{J}_c$ and the other due to the free current $\overline{J}_f$. Equation (15) can be rearranged to read

$$-\overline{E} \cdot \overline{J}_f^* = \nabla \cdot (\overline{E} \times \overline{H}^*) + \overline{E} \cdot \overline{J}_c^* + i\omega(\overline{E} \cdot \overline{D}^* - \overline{B} \cdot \overline{H}^*) \tag{17}$$

Consider a small volume element V. Equation (17) states that the complex power supplied to V by $\overline{J}_f$, $-\overline{E} \cdot \overline{J}_f^*$, is equal to the divergence of the complex Poynting power flow out of V, $\nabla \cdot (\overline{E} \times \overline{H}^*)$, plus the complex power dissipated in V, $\overline{E} \cdot \overline{J}_d^*$, plus the last term related to the stored complex electromagnetic energy in V.

e. Time-Average Poynting Power Vector

While the instantaneous value of the field vectors can be immediately determined from (2.1.1), the instantaneous value of the Poynting power density vector cannot be determined from the same rule; the power flow vector $\overline{S}$ involves the product of two field vectors. For more

insight into this issue, we let a complex field vector be represented by two real vectors. We write

$$\overline{E}(\bar{r}) = \overline{E}_R(\bar{r}) + i\overline{E}_I(\bar{r}) \tag{18}$$

where $\overline{E}_R$ and $\overline{E}_I$ are both real vectors representing the real and imaginary parts of the complex vector $\overline{E}$. Similarly,

$$\overline{H}(\bar{r}) = \overline{H}_R(\bar{r}) + i\overline{H}_I(\bar{r}) \tag{19}$$

The instantaneous values for the field vectors are

$$\overline{E}(\bar{r},t) = Re\{\overline{E}(\bar{r})e^{-i\omega t}\} = \overline{E}_R\cos\omega t + \overline{E}_I\sin\omega t \tag{20}$$

and

$$\overline{H}(\bar{r},t) = \overline{H}_R\cos\omega t + \overline{H}_I\sin\omega t \tag{21}$$

The complex Poynting's vector is

$$\overline{S} = \overline{E}\times\overline{H}^* = \overline{E}_R\times\overline{H}_R + \overline{E}_I\times\overline{H}_I + i(\overline{E}_I\times\overline{H}_R - \overline{E}_R\times\overline{H}_I) \tag{22}$$

We defined the instantaneous Poynting's vector $\overline{S}(\bar{r},t)$ as

$$\overline{S}(\bar{r},t) = \overline{E}(\bar{r},t)\times\overline{H}(\bar{r},t) \tag{23}$$

In view of (20) and (21), we have

$$\overline{S}(\bar{r},t) = \overline{E}_R\times\overline{H}_R\cos^2\omega t + \overline{E}_I\times\overline{H}_I\sin^2\omega t + (\overline{E}_R\times\overline{H}_I + \overline{E}_I\times\overline{H}_R)\sin\omega t\cos\omega t \tag{24}$$

Clearly (24) is not related to (22) in any way by the rule for field vectors as shown in (2.1.1). The instantaneous Poynting's vector $\overline{S}(\bar{r},t)$ is a real vector and is time-dependent. To relate $\overline{S}(\bar{r})$ to $\overline{S}(\bar{r},t)$ we must eliminate the time dependence in $\overline{S}(\bar{r},t)$. This is accomplished by a time averaging process. We find

$$\begin{aligned} <\overline{S}(\bar{r},t)> &= \frac{1}{2\pi}\int_0^{2\pi} d(\omega t)\,\overline{S}(\bar{r},t) \\ &= \frac{1}{2}\left[\overline{E}_R\times\overline{H}_R + \overline{E}_I\times\overline{H}_I\right] \\ &= \frac{1}{2}Re\left\{\overline{S}(\bar{r})\right\} \end{aligned} \tag{25}$$

where the first equality defines the time average of $\overline{S}(\overline{r},t)$, the second equality follows from (24), and the last equality follows from (22). Thus, when the complex Poynting's power vector $\overline{S} = \overline{E} \times \overline{H}^*$ is known, taking half of its real part yields the time average value of the instantaneous Poynting's vector:

$$< \overline{E} \times \overline{H} > = \frac{1}{2} Re \left\{ \overline{E} \times \overline{H}^* \right\} \tag{26}$$

This rule in general applies to the product of any two field quantities. That is, the time average of the product of two field quantities is equal to half of the real part of the product of one complex field quantity and the complex conjugate of the other complex field quantity.

f. Symmetry Conditions for Lossless Media

We now apply the complex Poynting's theorem to derive symmetry conditions for ***lossless media.*** Under time harmonic excitations, the constitutive matrices are usually complex and frequency-dependent. In source-free regions where $\overline{J} = 0$, the time average of the divergence of Poynting's vector, in view of the complex Poynting's theorem (15), is

$$< \nabla \cdot \overline{S} > = \frac{1}{2} Re \left\{ i\omega (\overline{H}^* \cdot \overline{B} - \overline{E} \cdot \overline{D}^*) \right\} \tag{27}$$

To find the real part of a complex number C, we note that $Re\{C\} = (C + C^*)/2$. Making use of this relation and substituting the constitutive relation for bianisotropic media in the $\overline{E}\,\overline{H}$ representation, we find

$$\begin{aligned} < \nabla \cdot \overline{S} > = \frac{i\omega}{4} \Big\{ & \overline{H}^* \cdot (\overline{\overline{\mu}} \cdot \overline{H} + \overline{\overline{\xi}} \cdot \overline{E}) - \overline{E} \cdot (\overline{\overline{\epsilon}}^* \cdot \overline{E}^* + \overline{\overline{\xi}}^* \cdot \overline{H}^*) \\ & - \overline{H} \cdot (\overline{\overline{\mu}}^* \cdot \overline{H}^* + \overline{\overline{\zeta}}^* \cdot \overline{E}^*) + \overline{E}^* \cdot (\overline{\overline{\epsilon}} \cdot \overline{E} + \overline{\overline{\xi}} \cdot \overline{H}) \Big\} \end{aligned}$$

For the term $\overline{H} \cdot \overline{\overline{\mu}}^* \cdot \overline{H}^*$, for instance, we make use of the relation $\overline{H} \cdot \overline{\overline{\mu}}^* \cdot \overline{H}^* = \overline{H}^* \cdot \overline{\overline{\mu}}^+ \cdot \overline{H}$ where the superscript + denotes the transpose and the complex conjugate of the matrix $\overline{\overline{\epsilon}}$, we find

$$\begin{aligned} < \nabla \cdot \overline{S} > = \frac{i\omega}{4} \Big\{ & \overline{E}^* \cdot (\overline{\overline{\epsilon}} - \overline{\overline{\epsilon}}^+) \cdot \overline{E} + \overline{H}^* \cdot (\overline{\overline{\mu}} - \overline{\overline{\mu}}^+) \cdot \overline{H} \\ & + \overline{E}^* \cdot (\overline{\overline{\xi}} - \overline{\overline{\zeta}}^+) \cdot \overline{H} + \overline{H}^* \cdot (\overline{\overline{\zeta}} - \overline{\overline{\xi}}^+) \cdot \overline{E} \Big\} \end{aligned} \tag{28}$$

The medium is characterized as passive if $< \nabla \cdot \overline{S} > < 0$, active if $< \nabla \cdot \overline{S} > > 0$, and lossless if $< \nabla \cdot \overline{S} >= 0$. For lossless media, (28) must vanish for all possible $\overline{E}$ and $\overline{H}$ fields, thus we obtain the lossless conditions

$$\overline{\overline{\epsilon}} = \overline{\overline{\epsilon}}^{+} \tag{29a}$$

$$\overline{\overline{\mu}} = \overline{\overline{\mu}}^{+} \tag{29b}$$

$$\overline{\overline{\xi}} = \overline{\overline{\zeta}}^{+} \tag{29c}$$

The lossless conditions (29*a*) and (29*b*) state that $\overline{\overline{\epsilon}}$ and $\overline{\overline{\mu}}$ are hermitian. Each contains six independent complex elements. Equation (29*c*) relates $\overline{\overline{\xi}}$ and $\overline{\overline{\zeta}}$; both matrices have a total of nine independent elements. Thus for a lossless bianisotropic medium, the constitutive relations contain a total of 21 independent complex elements.

The lossless conditions for the constitutive parameters in the $\overline{D}\,\overline{B}$ representations as shown in (1.1.9) can be derived to yield

$$\overline{\overline{\kappa}} = \overline{\overline{\kappa}}^{+}$$

$$\overline{\overline{\nu}} = \overline{\overline{\nu}}^{+}$$

$$\overline{\overline{\chi}} = \overline{\overline{\gamma}}^{+}$$

By the same token, we find

$$\overline{\overline{P}} = \overline{\overline{P}}^{+}$$

$$\overline{\overline{Q}} = \overline{\overline{Q}}^{+}$$

$$\overline{\overline{L}} = -\overline{\overline{M}}^{+}$$

for the constitutive parameters in the $\overline{E}\,\overline{B}$ representation.

2.2 Plane Wave Solutions

Consider source-free regions in a homogeneous isotropic medium characterized by a scalar permittivity ϵ and a scalar permeability μ. The Maxwell equations for the complex field vectors are

$$\nabla \times \overline{E} = i\omega\mu\overline{H} \tag{1}$$

$$\nabla \times \overline{H} = -i\omega\epsilon\overline{E} \tag{2}$$

$$\nabla \cdot \overline{E} = 0 \tag{3}$$

$$\nabla \cdot \overline{H} = 0 \tag{4}$$

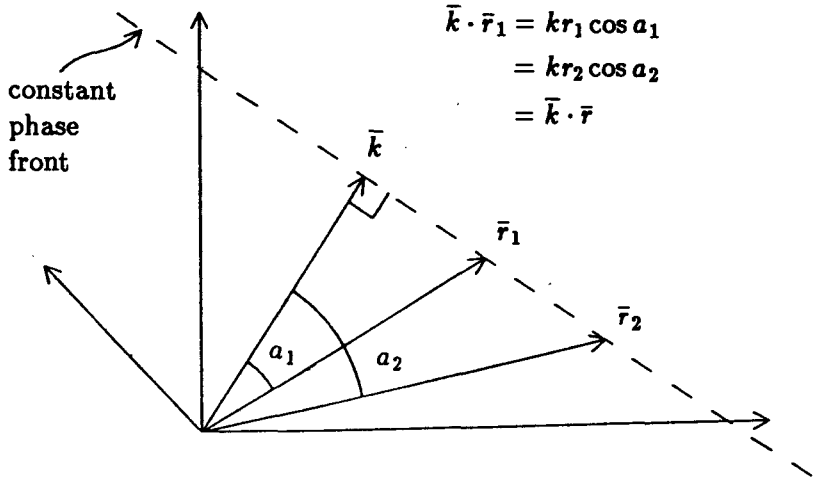


Figure 2.2.1 Constant phase front of a plane wave.

A wave equation for $\overline{E}$ can be readily derived by taking the curl of (1), introducing (2), and making use of (3). We obtain

$$(\nabla^2 + \omega^2 \mu\epsilon)\overline{E} = 0 \tag{5}$$

This is known as the homogeneous Helmholtz equation.

A general plane wave solution to (5) can be written as

$$\overline{E}(\bar{r}) = \overline{E}_0 \, e^{i\bar{k}\cdot\bar{r}} \tag{6}$$

The vector $\bar{k}$ is called the wave vector, the propagation vector, or simply the $\bar{k}$ vector. The exponential factor $\bar{k} \cdot \bar{r}$ can be written in the above solution in a rectangular coordinate system, as

$$\bar{k} \cdot \bar{r} = k_x x + k_y y + k_z z \tag{7}$$

We can write for the wave vector $\bar{k}$

$$\bar{k} = \hat{x}k_x + \hat{y}k_y + \hat{z}k_z \tag{8}$$

and for the position vector $\bar{r}$

$$\bar{r} = \hat{x}x + \hat{y}y + \hat{z}z \tag{9}$$

where $\hat{x}$, $\hat{y}$, and $\hat{z}$ denote unit vectors along the x, y, and z directions of the rectangular coordinate system, and k_x, k_y, and k_z are the components of the $\overline{k}$ vector.

For a given $\overline{k}$ vector, a constant phase front is determined by $\overline{k} \cdot \overline{r} =$ constant, which indicates that the front is perpendicular to the $\overline{k}$ vector [Fig. 2.2.1]. The phase front is a plane and the amplitude of the electric field on the plane is a constant. We call the solution in (6) a uniform plane wave. Since the constant phase front must be perpendicular to $\overline{k}$ at all times, we conclude that this phase front propagates in the direction of $\overline{k}$.

Substituting (6) in (5) yields the dispersion relation

$$k_x^2 + k_y^2 + k_z^2 = \omega^2 \mu\epsilon = k^2 \tag{10}$$

where k is the magnitude of the $\overline{k}$ vector. Substituting (6) in (1), we find the magnetic field vector

$$\overline{H} = \frac{1}{\omega\mu}\overline{k} \times \overline{E} = \frac{1}{\omega\mu}\overline{k} \times \overline{E}_0 \, e^{i\overline{k}\cdot\overline{r}}$$

The operation of the del operator ∇ on the spatial dependence $e^{i\overline{k}\cdot\overline{r}}$ is equivalent to replacing ∇ by $i\overline{k}$. The Maxwell equations (1)–(4) become, for the spatial dependence $e^{i\overline{k}\cdot\overline{r}}$ of the plane wave solutions,

$$\overline{k} \times \overline{E} = \omega\mu\overline{H} \tag{11}$$

$$\overline{k} \times \overline{H} = -\omega\epsilon\overline{E} \tag{12}$$

$$\overline{k} \cdot \overline{E} = 0 \tag{13}$$

$$\overline{k} \cdot \overline{H} = 0 \tag{14}$$

It is seen from (13) and (14) that $\overline{E}$ and $\overline{H}$ lie on the constant phase plane perpendicular to $\overline{k}$.

Considering the time evolution of $\overline{E}(r,t)$ at $r = 0$ by letting

$$\overline{E} = \overline{E}_R + i\overline{E}_I$$

we find

$$\overline{E}(t) = Re\{(\overline{E}_R + i\overline{E}_I)e^{-i\omega t}\} = \overline{E}_R \cos\omega t + \overline{E}_I \sin\omega t$$

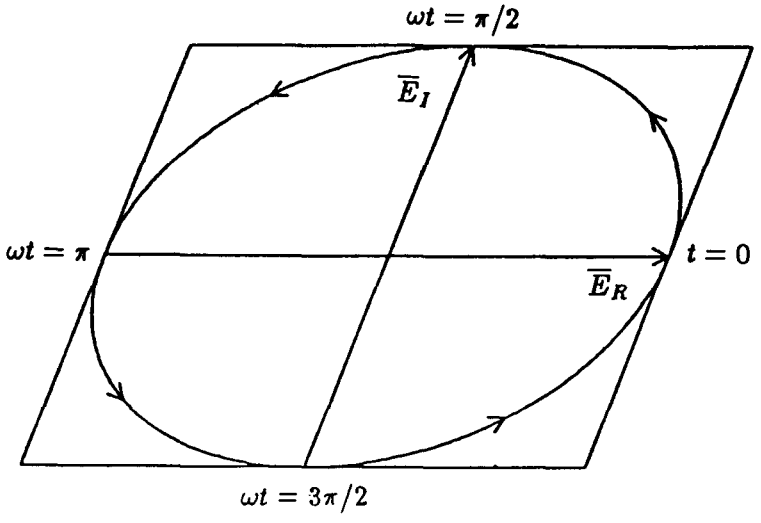


Figure 2.2.2 The polarization plane.

The time derivative of $\overline{E}(t)$ gives

$$\frac{\partial \overline{E}(t)}{\partial t} = -\omega[\overline{E}_R \sin \omega t - \overline{E}_I \cos \omega t]$$

In Figure 2.2.2 we sketch the two vectors $\overline{E}_R$ and $\overline{E}_I$. At $t = 0$, $\overline{E}(t)$ coincides with $\overline{E}_R$ and its time rate of change is parallel to $\overline{E}_I$. At $\omega t = \pi/2$, $\overline{E}(t)$ coincides with $\overline{E}_I$ and its time rate of change is in the opposite direction of $\overline{E}_R$. The tip of $\overline{E}$ as a function of time traces out an ellipse. The plane defined by the two vectors $\overline{E}_R$ and $\overline{E}_I$ is called the polarization plane. Thus the plane wave is in general elliptically polarized unless either $\overline{E}_R$ or $\overline{E}_I$ is zero; in this case the wave will be linearly polarized. For $\overline{k}$ in the direction pointing out of the paper, the motion of the tip of $\overline{E}(\overline{r}, t)$ follows the right-hand finger, while the thumb points in the propagation direction $\overline{k}$. The wave is right-hand elliptically polarized. When $\overline{E}_R$ and $\overline{E}_I$ are perpendicular to each other and have the same magnitudes, the wave will be circularly polarized.

a. Phase and Group Velocities

Consider the wave vector $\overline{k}$ given by (8). In the direction of $\hat{x}$, for instance, the plane perpendicular to $\hat{x}$ is not a constant phase

front. The phase distribution on this plane is given by $k_y y + k_z z$. The phase velocity and wavelength in the $\hat{x}$ direction are ω/k_x and $2\pi/k_x$, respectively. Thus the wave appears to have a larger phase velocity and longer wavelength in the $\hat{x}$ direction. The phase delay $\overline{T}_p$ and the group delay $\overline{T}_g$ are defined as

$$\overline{T}_p = \frac{\overline{k}}{\omega}$$

$$\overline{T}_g = \frac{d\overline{k}}{d\omega}$$

where $\overline{T}_g$ and $\overline{T}_p$ are both vectors. The larger the component of $\overline{k}$ in a particular direction, the larger the delay time for the phase front to travel a unit distance.

The phase and group velocities are defined as

$$v_{pi} = \frac{\omega}{k_i} \tag{15}$$

and

$$v_{gi} = \frac{d\omega}{dk_i} \tag{16}$$

where the subscript i denotes the ith component of the vector. Evidently the phase velocity vector v_{pi} reduces to $u = \omega/k$ only in the direction of the $\overline{k}$ vector. It is to be noted that a wave group can be formed from plane waves with different frequencies ω or from plane waves with different $\overline{k}$ vectors. Thus the difference in the phase and group velocities can be caused by media that are dispersive or non-isotropic. When a medium is not isotropic, the magnitude of $\overline{k}$ may vary as a function of its direction.

Because the dispersion relation pertaining to the medium provides a relation for ω and $\overline{k}$, it is necessary only to specify two components of the $\overline{k}$ vector in order to determine a wave motion. Equivalently, a specification of the direction of $\overline{k}$ is also sufficient. The magnitude of $\overline{k}$ is then determined from the dispersion relation in (10). When the medium is not isotropic, the magnitude of $\overline{k}$ also varies as a function of its direction.

Consider a dispersive medium such as the electron plasma with

$$\epsilon = \epsilon_o \left[1 - \frac{\omega_p^2}{\omega^2}\right]$$

The dispersion relation $k = \omega(\mu_o\epsilon)^{1/2}$ is then a nonlinear function of ω. We write the wavenumber $k = k(\omega)$. Waves of different frequencies propagate at different phase velocities. Consider a group of plane waves with different angular frequencies propagating along the $\hat{z}$ direction in a dispersive medium. Assume that all angular frequencies are in the neighborhood of a center frequency ω_0. We can express the wavenumber k around ω_0,

$$k(\omega) = k(\omega_0) + (\omega - \omega_0)\left[\frac{\delta k(\omega)}{\delta\omega}\right]_{\omega=\omega_0} + \frac{1}{2}(\omega - \omega_0)^2\left[\frac{\delta^2 k(\omega)}{\delta\omega^2}\right]_{\omega=\omega_0} + \cdots \tag{17}$$

For a narrow-band signal, we can retain only the first two terms. The space–time dependence of the group of plane waves becomes

$$e^{ik(\omega)z - i\omega t} = e^{ik(\omega_0)z - i\omega_0 t} e^{-i(\omega-\omega_0)\left\{t - z\left[\frac{dk(\omega)}{d\omega}\right]_{\omega=\omega_0}\right\}}$$

which can be viewed as a wave propagating with phase delay $T_p = k/\omega_0$ and group delay $T_g = dk(\omega)/d\omega$.

We can construct a $\omega - k$ diagram, plotting ω as a function of k. Figure 2.2.3 shows the $\omega - k$ diagram for an isotropic plasma medium. The slope of the straight line from the origin to a point on the curve at $\omega = \omega_0$ then represents the phase velocity v_p, and the slope at $\omega = \omega_0$ represents the group velocity v_g. We find

$$v_p = \frac{1}{\sqrt{\mu_o\epsilon_o\left(1 - \omega_p^2/\omega^2\right)}}$$

and

$$v_g = \frac{\sqrt{1 - \omega_p^2/\omega^2}}{\sqrt{\mu_o\epsilon_o}}$$

It is seen that v_p is larger than the velocity of light $c = (\mu_o\epsilon_o)^{-1/2}$ and that $v_p v_g = c^2$. For a nondispersive medium, the $\omega - k$ curves are straight lines starting from the origin, and the phase delay and

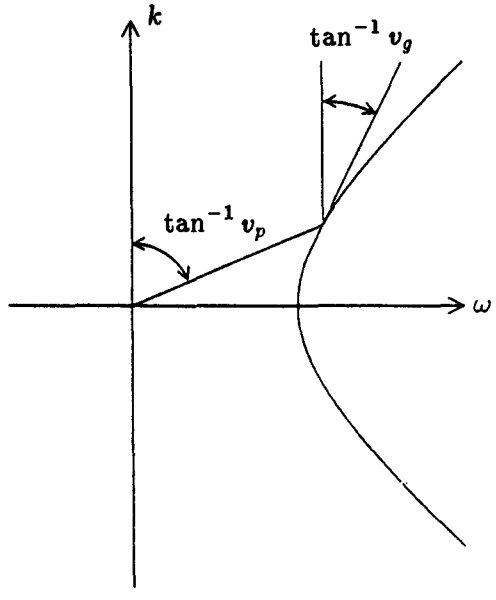


Figure 2.2.3 $\omega - k$ **diagram for plasma medium.**

the group delay are the same, as are the phase velocity and the group velocity.

b. Evanescent Waves in Lossless Media

To illustrate the use of Poynting's power density vector $\overline{S}$, we first show that a plane wave may be attenuated in a lossless medium but loses no power. As an example, consider the isotropic plasma with $\epsilon = \epsilon_o(1 - \omega_p^2/\omega^2)$. For $\omega < \omega_p$, $k = \omega(\mu\epsilon)^{1/2} = ik_I$ with $k_I = \omega[\mu\epsilon_o(\omega_p^2/\omega^2 - 1)]^{1/2}$. For a plane wave propagating along the $\hat{z}$ direction with $\overline{E}$ in the $\hat{x}$ direction, we have

$$\overline{E} = \hat{x}E_0\, e^{-k_I z} \tag{18}$$

$$\overline{H} = \hat{y}\frac{ik_I}{\omega\mu}E_0\, e^{-k_I z} \tag{19}$$

The complex Poynting power density

$$\overline{S} = \overline{E} \times \overline{H}^* = -\hat{z}\frac{ik_I}{\omega\mu}\,|E_0|^2\, e^{-2k_I z} \tag{20}$$

The time-average Poynting power density

$$< \overline{S} >= \frac{1}{2} Re\{\overline{S}\} = 0 \tag{21}$$

Thus the wave amplitude decay with an attenuation constant k_I but no power is dissipated in the medium. A wave that attenuates exponentially in a medium and dissipates no power is called an *evanescent wave.*

c. *Penetration Depth in Lossy Media*

For lossy media, $\epsilon = \epsilon_R + i\epsilon_I$ is complex. The wavenumber $k = \omega(\mu\epsilon)^{1/2}$ is complex too. A wave propagating in a certain direction also decays exponentially in that direction. Let $k = k_R + ik_I$, then in a distance of $1/k_I$ along the direction of propagation, the wave amplitude will decay by a factor of e^{-1}. We call

$$d_p = \frac{1}{k_I} \tag{22}$$

the *penetration depth.*

Consider conducting media with permittivity $\epsilon + i\sigma/\omega$. For a highly conducting medium with $1 \ll \sigma/\omega\epsilon$, we approximate

$$k \approx \omega(\mu\epsilon)^{1/2} \left[i\frac{\sigma}{\omega\epsilon}\right]^{1/2} = (1+i)\sqrt{\frac{\omega\mu\sigma}{2}}$$

We find the penetration depth

$$d_p = \sqrt{\frac{2}{\omega\mu\sigma}} = \delta \tag{23}$$

which is usually a very small number also known as the *skin depth.*

For a slightly conducting medium with $\sigma/\omega\epsilon \ll 1$, we can approximate

$$k = \omega(\mu\epsilon)^{1/2}\left[1 + \frac{i\sigma}{\omega\epsilon}\right]^{1/2} \approx \omega(\mu\epsilon)^{1/2}\left[1 + \frac{i\sigma}{2\omega\epsilon}\right] = \omega(\mu\epsilon)^{1/2} + i\frac{\sigma}{2}\left(\frac{\mu}{\epsilon}\right)^{1/2}$$

We find the penetration depth

$$d_p = \frac{2}{\sigma}\left(\frac{\epsilon}{\mu}\right)^{1/2} \tag{24}$$

It is interesting to note that the penetration depth in (24) is independent of frequency. Physically we expect that there will be less penetration at higher frequencies. However we have assumed a homogeneous medium here, and high frequency attenuation is largely due to scattering, which is the subject matter of Chapter VI.

In a lossy medium, power is dissipated. Let a plane wave be propagating along the $\hat{z}$ direction with electric field $\overline{E}$ in the $\hat{x}$ direction,

$$\overline{E} = \hat{x} E_0 \, e^{i(k_R + ik_I)z} \tag{25}$$

The corresponding magnetic field is

$$\overline{H} = \frac{1}{i\omega\mu} \nabla \times \overline{E} = \hat{y} \frac{k_R + ik_I}{\omega\mu} E_0 \, e^{i(k_R + ik_I)z} \tag{26}$$

The time-average Poynting power density propagating in the medium is found to be

$$< \overline{S} > = \frac{1}{2} Re \left\{ \overline{E} \times \overline{H}^* \right\} = \hat{z} \frac{1}{2} \frac{k_R}{\omega\mu} |E_0|^2 \, e^{-2k_I z} \tag{27}$$

The power density attenuates by an amount of $1/e^2$ in a penetration depth. This attenuation accounts for the power dissipated in the lossy medium.

The components of a $\overline{k}$ vector may not all be real, even in a lossless medium where the magnitude of k is real. Consider an isotropic medium with the dispersion relation

$$k^2 = k_x^2 + k_y^2 + k_z^2 = \omega^2 \mu\epsilon$$

We see that the components of the $\overline{k}$ vector, k_x, k_y, and k_z, can have complex values even for real ϵ and μ. The imaginary part of a complex component in a particular direction indicates that the wave either grows or decays exponentially in that direction. Let the $\overline{k}$ vector of a plane wave possess two components k_y and k_x, such that k_y is imaginary, $k_y = ik_{yI}$, and $k_x^2 - k_{yI}^2 = \omega^2\mu\epsilon$. The wave then behaves as $\exp(ik_x x) \exp(-k_{yI} y)$. It propagates in the $\hat{x}$ direction with phase delay k_x/ω and decays exponentially in the $\hat{y}$ direction. We call the wave *evanescent* in the $\hat{y}$ direction. It can also be called a *nonuniform plane wave* propagating in the $\hat{x}$ direction because the constant phase front, which is a plane perpendicular to the $\hat{x}$ direction, no longer has

uniform wave amplitudes. The evanescent nonuniform plane wave will be discussed in Chapter III.

2.3 Plane Waves in Homogeneous Media and the *kDB* System

We shall now investigate solutions to the Maxwell equations in regions void of source, namely in regions where $\overline{J} = \rho = 0$. This, of course, does not mean that there is no source anywhere in space. Sources must exist outside the regions of interest in order to produce fields in these regions. Thus, in source-free regions, Maxwell's equations for a harmonic field become

$$\nabla \times \overline{E} = i\omega \overline{B} \tag{1}$$

$$\nabla \times \overline{H} = -i\omega \overline{D} \tag{2}$$

$$\nabla \cdot \overline{B} = 0 \tag{3}$$

$$\nabla \cdot \overline{D} = 0 \tag{4}$$

We also assume homogeneous media where constitutive relations are independent of spatial coordinates. Plane wave solutions of the form $e^{i\overline{k}\cdot\overline{r}}$ are then admissible. Letting all complex field vectors have the same spatial dependence $e^{i\overline{k}\cdot\overline{r}}$, we obtain from (1)–(4)

$$\overline{k} \times \overline{E} = \omega \overline{B} \tag{5}$$

$$\overline{k} \times \overline{H} = -\omega \overline{D} \tag{6}$$

$$\overline{k} \cdot \overline{B} = 0 \tag{7}$$

$$\overline{k} \cdot \overline{D} = 0 \tag{8}$$

We see from (7) and (8) that the complex vectors $\overline{D}$ and $\overline{B}$ are always perpendicular to the wave vector $\overline{k}$. We call this plane, which is perpendicular to $\overline{k}$ and contains both $\overline{D}$ and $\overline{B}$, the DB plane. For a medium with $\overline{B} = \mu \overline{H}$, the vector $\overline{H}$ also lies in the DB plane. If the medium is anisotropic with $\overline{D} = \overline{\overline{\epsilon}} \cdot \overline{E}$, we see that $\overline{E}$ may not lie on the DB plane. For this reason we shall define the polarization of the plane wave by using the vector $\overline{D}$ instead of $\overline{E}$ as we did in previous sections. Notice that the Poynting vector is in the direction of $\overline{E} \times \overline{H}$, which is not necessarily in the same direction of $\overline{k}$ inside an anisotropic medium. Thus the direction of power flow of a plane wave

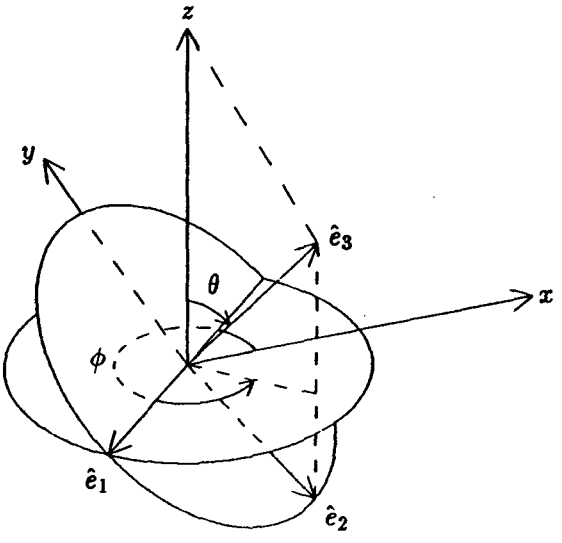


Figure 2.3.1 The kDB system.

may not always be in the direction of the wave vector $\overline{k}$. We shall explore in detail the propagation characteristics in several anisotropic media.

a. kDB System

To facilitate discussions on wave behavior and solutions for the field vectors inside a general homogeneous medium, we first establish a convenient coordinate system called the kDB system, which consists of the $\overline{k}$ vector and the DB plane. The kDB system has unit vectors $\hat{e}_1$, $\hat{e}_2$, and $\hat{e}_3$. We let $\hat{e}_3$ be in the direction of $\overline{k}$ so that $\overline{k} = \hat{e}_3 k$. In terms of the xyz coordinate system [Fig. 2.3.1], we find

$$\hat{e}_3 = \hat{x}\sin\theta\cos\phi + \hat{y}\sin\theta\sin\phi + \hat{z}\cos\theta \tag{9}$$

The unit vector $\hat{e}_1$ lies in the $x-y$ plane and is perpendicular to the projection of $\overline{k}$ on the $x-y$ plane. When $\theta \neq 0$ it is determined by the intersection of the $x-y$ plane and the DB plane. We let

$$\hat{e}_1 = \hat{x}\sin\phi - \hat{y}\cos\phi \tag{10}$$

which is seen to be perpendicular to $\hat{e}_3$ because $\hat{e}_1 \cdot \hat{e}_3 = 0$. The unit vector $\hat{e}_2$ is easily determined by requiring that $\hat{e}_2$ be perpendicular

to both $\hat{e}_1$ and $\hat{e}_3$ and follows the right-hand rule. We let

$$\hat{e}_2 = \hat{e}_3 \times \hat{e}_1 = \hat{x}\cos\theta\cos\phi + \hat{y}\cos\theta\sin\phi - \hat{z}\sin\theta \tag{11}$$

The three unit vectors are indicated in Figure 2.3.1. We see that if the $x-y$ plane is rotated counter-clockwise about $\hat{z}$ by $\phi-\pi/2$, and then rotated about the new x axis by θ, the resulting plane perpendicular to $\overline{k}$ is the DB plane.

We now establish transformation formulas for components of field vectors. Let a vector be called $\overline{A}$ when it is represented by components projected onto the xyz coordinate system. We write

$$\overline{A} = \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} \tag{12}$$

Let the same vector be called $\overline{A}_k$ when it is represented by components projected onto the kDB coordinate system,

$$\overline{A}_k = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} \tag{13}$$

The relation between the components of $\overline{A}$ and the components of $\overline{A}_k$ are governed by

$$\overline{A}_k = \overline{\overline{T}} \cdot \overline{A} \tag{14}$$

or

$$\overline{A} = \overline{\overline{T}}^{-1} \cdot \overline{A}_k \tag{15}$$

where $\overline{\overline{T}}^{-1}$ is the inverse of $\overline{\overline{T}}$. We wish to determine the transformation matrix $\overline{\overline{T}}$ and its inverse $\overline{\overline{T}}^{-1}$.

Since $\overline{A}$ and $\overline{A}_k$ denote the same vector, we find

$$\begin{aligned} A_1 &= \hat{e}_1 \cdot A = \hat{e}_1 \cdot \hat{x}A_x + \hat{e}_1 \cdot \hat{y}A_y + \hat{e}_1 \cdot \hat{z}A_z \\ &= \sin\phi A_x - \cos\phi A_y \end{aligned} \tag{16}$$

$$\begin{aligned} A_2 &= \hat{e}_2 \cdot A = \hat{e}_2 \cdot \hat{x}A_x + \hat{e}_2 \cdot \hat{y}A_y + \hat{e}_2 \cdot \hat{z}A_z \\ &= \cos\theta\cos\phi A_x + \cos\theta\sin\phi A_y - \sin\theta A_z \end{aligned} \tag{17}$$

$$\begin{aligned} A_3 &= \hat{e}_3 \cdot \overline{A} = \hat{e}_3 \cdot \hat{x}A_x + \hat{e}_3 \cdot \hat{y}A_y + \hat{e}_3 \cdot \hat{z}A_z \\ &= \sin\theta\cos\phi A_x + \sin\theta\sin\phi A_y + \cos\theta A_z \end{aligned} \tag{18}$$

where we have made use of (9)–(11). Writing (16)–(18) in matrix form and comparing with (14), we obtain

$$\overline{\overline{T}} = \begin{pmatrix} \sin\phi & -\cos\phi & 0 \\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{pmatrix} \tag{19}$$

The inverse of $\overline{\overline{T}}$ is simply calculated as

$$\overline{\overline{T}}^{-1} = \begin{pmatrix} \sin\phi & \cos\theta\cos\phi & \sin\theta\cos\phi \\ -\cos\phi & \cos\theta\sin\phi & \sin\theta\sin\phi \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \tag{20}$$

which is seen to be the transpose of $\overline{\overline{T}}$. The result (20) can be obtained in three different ways: (i) by directly calculating the inverse of $\overline{\overline{T}}$; (ii) by following a procedure simliar to (16)–(18) and using instead the relations $A_x = \hat{x}\cdot\overline{A}_k$, $A_y = \hat{y}\cdot\overline{A}_k$, and $A_z = \hat{z}\cdot\overline{A}_k$; and (iii) by recognizing that the transformation is orthogonal so that the inverse of $\overline{\overline{T}}$ is equal to the transpose of $\overline{\overline{T}}$. Clearly the product of (19) and (20) is a unit matrix.

The transformation formula established in (14)–(15) applies to any field vector $\overline{E}$, $\overline{D}$, $\overline{B}$, or $\overline{H}$. We must now find formulas that transform constitutive relations from xyz to kDB. Notice the fact that vectors $\overline{D}$ and $\overline{B}$ take much simpler forms than $\overline{E}$ and $\overline{H}$ because $D_3 = B_3 = 0$. We express the constitutive relations in xyz by relating $\overline{E}$ and $\overline{H}$ to $\overline{B}$ and $\overline{D}$,

$$\overline{E} = \overline{\overline{\kappa}}\cdot\overline{D} + \overline{\overline{\chi}}\cdot\overline{B} \tag{21}$$

$$\overline{H} = \overline{\overline{\nu}}\cdot\overline{B} + \overline{\overline{\gamma}}\cdot\overline{D} \tag{22}$$

We call these the constitutive relations in the $\overline{D}\,\overline{B}$ representation. Making use of the transformation formula (15) we find $\overline{E} = \overline{\overline{T}}^{-1}\cdot\overline{E}_k$ and similarly for $\overline{H}$, $\overline{D}$ and $\overline{B}$. The result is

$$\overline{E}_k = (\overline{\overline{T}}\cdot\overline{\overline{\kappa}}\cdot\overline{\overline{T}}^{-1})\cdot\overline{D}_k + (\overline{\overline{T}}\cdot\overline{\overline{\chi}}\cdot\overline{\overline{T}}^{-1})\cdot\overline{B}_k \tag{23}$$

$$\overline{H}_k = (\overline{\overline{T}}\cdot\overline{\overline{\nu}}\cdot\overline{\overline{T}}^{-1})\cdot\overline{B}_k + (\overline{\overline{T}}\cdot\overline{\overline{\gamma}}\cdot\overline{\overline{T}}^{-1})\cdot\overline{D}_k \tag{24}$$

We thus obtain

$$\overline{\overline{\kappa}}_k = \overline{\overline{T}}\cdot\overline{\overline{\kappa}}\cdot\overline{\overline{T}}^{-1} \tag{25}$$

$$\overline{\overline{\chi}}_k = \overline{\overline{T}} \cdot \overline{\overline{\chi}} \cdot \overline{\overline{T}}^{-1} \tag{26}$$

$$\overline{\overline{\nu}}_k = \overline{\overline{T}} \cdot \overline{\overline{\nu}} \cdot \overline{\overline{T}}^{-1} \tag{27}$$

$$\overline{\overline{\gamma}}_k = \overline{\overline{T}} \cdot \overline{\overline{\gamma}} \cdot \overline{\overline{T}}^{-1} \tag{28}$$

such that in the kDB system

$$\overline{E}_k = \overline{\overline{\kappa}}_k \cdot \overline{D}_k + \overline{\overline{\chi}}_k \cdot \overline{B}_k \tag{29}$$

$$\overline{H}_k = \overline{\overline{\nu}}_k \cdot \overline{B}_k + \overline{\overline{\gamma}}_k \cdot \overline{D}_k \tag{30}$$

With the transformation formulas (14) and (25)–(28), we can now transform all quantities from the xyz coordinate system to the kDB system.

b. Maxwell's Equations in kDB System

Within the frame of the kDB system, the Maxwell equations for plane waves inside source-free homogeneous media take the same form as (5)–(8)

$$\overline{k} \times \overline{E}_k = \omega \overline{B}_k \tag{31}$$

$$\overline{k} \times \overline{H}_k = -\omega \overline{D}_k \tag{32}$$

$$\overline{k} \cdot \overline{B}_k = 0 \tag{33}$$

$$\overline{k} \cdot \overline{D}_k = 0 \tag{34}$$

but with the $\overline{k}$ vector in the $\hat{e}_3$ direction

$$\overline{k} = \hat{e}_3 k \tag{35}$$

From (33)–(34) we know that $B_3 = D_3 = 0$. Making use of the constitutive relations (29)–(30), we find from (31)–(32)

$$\omega B_2 = kE_1 = k(\kappa_{11} D_1 + \kappa_{12} D_2 + \chi_{11} B_1 + \chi_{12} B_2)$$

$$\omega B_1 = -kE_2 = -k(\kappa_{21} D_1 + \kappa_{22} D_2 + \chi_{21} B_1 + \chi_{22} B_2)$$

$$\omega D_2 = -kH_1 = -k(\nu_{11} B_1 + \nu_{12} B_2 + \gamma_{11} D_1 + \gamma_{12} D_2)$$

$$\omega D_1 = kH_2 = k(\nu_{21} B_1 + \nu_{22} B_2 + \gamma_{21} D_1 + \gamma_{22} D_2)$$

We first divide both sides by k and let $u = \omega/k$. Rearranging terms and writing in matrix form, we obtain

$$\begin{pmatrix} \kappa_{11} & \kappa_{12} \\ \kappa_{21} & \kappa_{22} \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = - \begin{pmatrix} \chi_{11} & \chi_{12} - u \\ \chi_{21} + u & \chi_{22} \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \tag{36}$$

$$\begin{pmatrix} \nu_{11} & \nu_{12} \\ \nu_{21} & \nu_{22} \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = - \begin{pmatrix} \gamma_{11} & \gamma_{12} + u \\ \gamma_{21} - u & \gamma_{22} \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} \tag{37}$$

We can further eliminate either $\overline{B}_k$ or $\overline{D}_k$ from (36) and (37) and derive a 2×2 matrix equation for $\overline{D}_k$ or $\overline{B}_k$ alone. The corresponding two equations are linear and homogeneous. Setting the determinant of the 2×2 matrix equal to zero, we then obtain the dispersion relations for the homogeneous media. In the following sections we shall discuss the applications of the kDB system in the study of plane wave characteristics in homogeneous media.

We notice that the procedure for deriving dispersion relations with the kDB system is also applicable to dissipative media, where $\overline{k} = \hat{x}k_x + \hat{y}k_y + \hat{z}k_z$ is a complex vector. We first treat k and the angles θ and ϕ as though they were real. After the solution is obtained, we eliminate θ and ϕ by relating them to the rectangular components of $\overline{k}$ in the original xyz coordinate system and allowing the $\overline{k}$ vector to be complex.

As an example consider isotropic media for which the constitutive relations in $\overline{D}\,\overline{B}$ representation are

$$\overline{E} = \kappa \overline{D}$$
$$\overline{H} = \nu \overline{B}$$

where $\kappa = 1/\epsilon$ is called the impermittivity and $\nu = 1/\mu$ is called the impermeability.

In the kDB system, we find that

$$\overline{E}_k = \kappa \overline{D}_k$$
$$\overline{H}_k = \nu \overline{B}_k$$

Substituting in (36)–(37), and recognizing that $\kappa_{11} = \kappa_{22} = \kappa$, $\nu_{11} = \nu_{22} = \nu$ and that all other constitutive elements are equal to zero, we have

$$\kappa \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = \begin{pmatrix} 0 & u \\ -u & 0 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$$

$$\nu \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} 0 & -u \\ u & 0 \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix}$$

Eliminating $\overline{B}_k$ from the above two equations yields

$$\begin{pmatrix} u^2 - \kappa\nu & 0 \\ 0 & u^2 - \kappa\nu \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = 0$$

We see that when there is a wave field in the media, the following relation must be satisfied:

$$u^2 - \kappa\nu = 0$$

which is the dispersion relation for isotropic media.

The phase velocity $u = \omega/k$ of the plane wave is

$$u = \pm\sqrt{\kappa\nu} = \pm\frac{1}{\sqrt{\mu\epsilon}}$$

The $\pm$ signs refer to waves propagating in opposite directions. Written in terms of ω and k, the dispersion relation becomes

$$k^2 = \omega^2\mu\epsilon$$

When the medium is dispersive, that is, when μ or ϵ is a function of ω or k, the dispersion relation can be a very complicated function relating k and ω.

The dispersion relations for plane waves in homogeneous media can be derived in a number of different ways. For instance, making use of the constitutive relations in $\overline{E}\,\overline{H}$ representation

$$\overline{D} = \overline{\overline{\epsilon}} \cdot \overline{E} + \overline{\overline{\xi}} \cdot \overline{H}$$
$$\overline{B} = \overline{\overline{\mu}} \cdot \overline{H} + \overline{\overline{\zeta}} \cdot \overline{E}$$

we can eliminate the field vectors $\overline{D}$, $\overline{B}$ and $\overline{H}$ from the above constitutive relations and (5)–(6). Defining an operator $\overline{\overline{k}}$ such that $\overline{\overline{k}} \cdot \overline{A} = \overline{k} \times \overline{A}$ for any vector $\overline{A}$, we obtain

$$\left\{\omega^2\overline{\overline{\epsilon}} + [\overline{\overline{k}} + \omega\overline{\overline{\xi}}] \cdot \overline{\overline{\mu}}^{-1} \cdot [\overline{\overline{k}} - \omega\overline{\overline{\zeta}}]\right\} \cdot \overline{E} = 0$$

For nontrivial solutions of $\overline{E}$, the determinant of the matrix operating on $\overline{E}$ must be equal to zero. Hence

$$\left|\omega^2\overline{\overline{\epsilon}} + [\overline{\overline{k}} + \omega\overline{\overline{\xi}}] \cdot \overline{\overline{\mu}}^{-1} \cdot [\overline{\overline{k}} - \omega\overline{\overline{\zeta}}]\right| = 0$$

This is the dispersion relation relating the components of $\overline{k}$ and the angular frequency ω. It reduces to the isotropic case in a straightforward manner. However, in the cases of non-isotropic media, the study of the wave behavior becomes extremely involved. It is the kDB system that provides a systematic approach in facilitating the interpretation of the various plane wave characteristics in more general media.

2.4 Plane Waves in Uniaxial Media

In the principal, xyz, coordinate system of a uniaxial medium the constitutive relations under $\overline{D}\,\overline{B}$ representation are

$$\overline{E} = \overline{\overline{\kappa}} \cdot \overline{D} \tag{1}$$

$$\overline{H} = \nu \overline{B} \tag{2}$$

where

$$\overline{\overline{\kappa}} = \begin{pmatrix} \kappa & 0 & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & \kappa_z \end{pmatrix} \tag{3}$$

is called the impermittivity tensor. The optic axis is in the $\hat{z}$ direction. In terms of the permittivity tensor

$$\overline{\overline{\epsilon}} = \begin{pmatrix} \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon_z \end{pmatrix} \tag{4}$$

we find that $\kappa = 1/\epsilon$ and $\kappa_z = 1/\epsilon_z$, since $\overline{\overline{\kappa}} = \overline{\overline{\epsilon}}^{-1}$. The impermeability is related to the permeablity μ by $\nu = 1/\mu$. Transforming to the kDB system, we find from (2.3.25), that

$$\overline{\overline{\kappa}}_k = \overline{\overline{T}} \cdot \overline{\overline{\kappa}} \cdot \overline{\overline{T}}^{-1} = \begin{pmatrix} \kappa & 0 & 0 \\ 0 & \kappa\cos^2\theta + \kappa_z \sin^2\theta & (\kappa - \kappa_z)\sin\theta\cos\theta \\ 0 & (\kappa - \kappa_z)\sin\theta\cos\theta & \kappa\sin^2\theta + \kappa_z\cos^2\theta \end{pmatrix} \tag{5}$$

Since the uniaxial medium has cylindrical symmetry around the z axis, the transformed relation is ϕ-independent as one should expect. Applying (2.3.36)–(2.3.37), we obtain

$$\begin{pmatrix} \kappa_{11} & 0 \\ 0 & \kappa_{22} \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = \begin{pmatrix} 0 & u \\ -u & 0 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \tag{6}$$

$$\nu \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} 0 & -u \\ u & 0 \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} \tag{7}$$

Note that $\overline{\overline{\chi}} = \overline{\overline{\gamma}} = 0$ and, from (5), $\kappa_{11} = \kappa$ and $\kappa_{22} = \kappa\cos^2\theta + \kappa_z \sin^2\theta$. Note also that although κ_{23} and κ_{33} are both calculated in (5), they will not be needed either in applying (6) or in determining $\overline{E}_k$ from $\overline{D}_k$, as illustrated later. Eliminating $\overline{B}_k$ from (6) and (7),

$$\begin{pmatrix} u^2 - \nu\kappa_{11} & 0 \\ 0 & u^2 - \nu\kappa_{22} \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = 0 \tag{8}$$

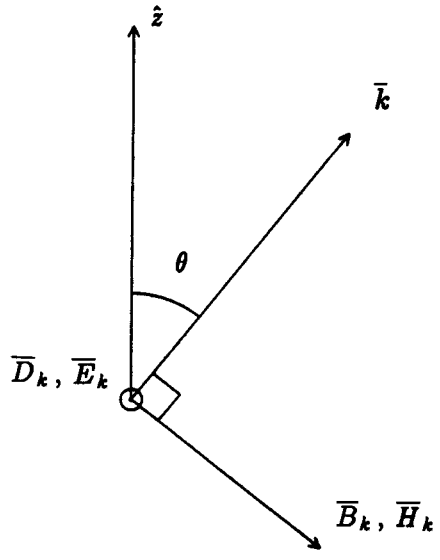


Figure 2.4.1 Ordinary wave in uniaxial medium.

In order to satisfy (8), we have the following four cases:

(A) $D_1 = D_2 = 0$
(B) $D_1 \neq 0$ and $D_2 = 0$, $\quad u^2 - \nu\kappa_{11} = 0$
(C) $D_1 = 0$ and $D_2 \neq 0$, $\quad u^2 - \nu\kappa_{22} = 0$
(D) $D_1 \neq 0$ and $D_2 \neq 0$, $\quad u^2 - \nu\kappa_{11} = u^2 - \nu\kappa_{22} = 0$

Case A implies that there is no field at all.

a. Ordinary and Extraordinary Waves

Case B corresponds to a wave which is linearly polarized in the $\hat{e}_1$ direction. Notice from Figure 2.4.1 that $\hat{e}_1$ is perpendicular to the plane formed by the optic axis and the $\overline{k}$ vector. This linearly polarized plane wave propagates with the phase velocity

$$u = \pm\sqrt{\nu\kappa_{11}} \tag{9}$$

The other field components for the wave are determined from (7) and the constitutive relations (1) and (2). We find that

$$\overline{D}_k = \hat{e}_1 D_1 \tag{10}$$

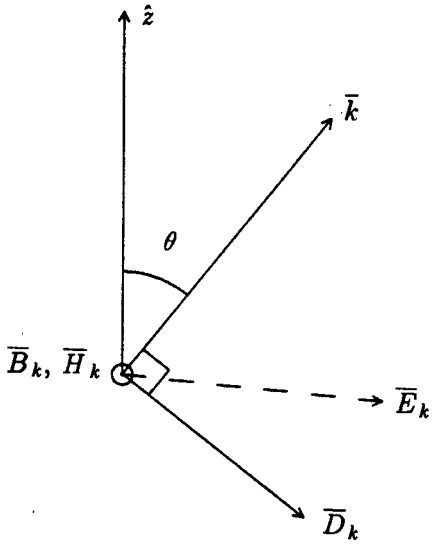


Figure 2.4.2 Extraordinary wave in uniaxial medium.

$$\overline{B}_k = \hat{e}_2 \frac{u}{\nu} D_1 \tag{11}$$

$$\overline{H}_k = \hat{e}_2 \, u D_1 \tag{12}$$

$$\overline{E}_k = \hat{e}_1 \, \kappa D_1 \tag{13}$$

Thus $\overline{D}_k$ and $\overline{E}_k$ are in the same direction, as are $\overline{B}_k$ and $\overline{H}_k$ [Fig. 2.4.1]. We call this the ordinary wave in the uniaxial medium.

For case C, the wave is linearly polarized in the $\hat{e}_2$ direction. Remember that $\hat{e}_2$ lies in the plane determined by the optic axis and the $\overline{k}$ vector and is perpendicular to the $\overline{k}$ vector. This linearly polarized wave propagates with the phase velocity

$$u = \pm\sqrt{\nu\kappa_{22}} = \pm[\nu(\kappa \cos^2\theta + \kappa_z \sin^2\theta)]^{1/2} \tag{14}$$

which is seen to be dependent upon the direction of propagation. The other field components for the wave are determined from (7) and the constitutive relations (1) and (2). We find that

$$\overline{D}_k = \hat{e}_2 \, D_2 \tag{15}$$

$$\overline{B}_k = -\hat{e}_1 \frac{u}{\nu} D_2 \tag{16}$$

$$\overline{H}_k = -\hat{e}_1 \, u D_2 \tag{17}$$

$$\overline{E}_k = \hat{e}_2 \, \kappa_{22} D_2 + \hat{e}_3 \, (\kappa - \kappa_z) \sin\theta \cos\theta \, D_2 \tag{18}$$

We see that $\overline{E}_k$ and $\overline{D}_k$ both lie in the plane determined by the optic axis $\hat{z}$ and the wave vector $\overline{k}$ but are no longer in the same direction. For a positive uniaxial medium, $\epsilon_z > \epsilon$, $\overline{E}_k$ lies between the vectors $\overline{k}$ and $\overline{D}_k$ [Fig. 2.4.2]. For a negative uniaxial medium, $\overline{E}_k$ makes an angle larger than $\pi/2$ with $\overline{k}$. As a consequence, the direction of the Poynting's vector will not be in the direction parallel to $\overline{k} = \hat{e}_3 k$. This is seen from the cross-product of $\overline{E}_k \times \overline{H}_k^*$. This linearly polarized wave, the phase velocity of which has magnitude dependent on angle and direction different from that of the Poynting's vector, is called the extraordinary wave in the uniaxal medium.

In order for case D to apply, that is, $D_1 \neq 0$ and $D_2 \neq 0$, we must have $\kappa_{11} = \kappa_{22}$, which cannot hold unless (i) the medium is isotropic or (ii) the direction of propagation is along $\hat{z}$. Thus waves with polarizations other than linear can propagate only when the $\overline{k}$ vector is along the direction of the optic axis. In general, plane waves inside a uniaxial medium are either an ordinary wave linearly polarized with the $\overline{D}$ vector perpendicular to the plane determined by the optic axis and the $\overline{k}$ vector, propagating with the phase velocity in (9), or an extraordinary wave linearly polarized with the $\overline{D}$ vector lying in the plane of the optic axis and the $\overline{k}$ vector and perpendicular to the $\overline{k}$ vector, propagating with the phase velocity in (14). The result of these two characteristic waves propagating with different phase velocities in a medium is called birefringence and the medium is called birefringent. When an electromagnetic wave enters a uniaxial medium, it decomposes into the two linearly polarized characteristic waves which propagate with different velocities. This phenomemon is known as double refraction. Consider a slab of uniaxial medium with the y axis perpendicular to its front and back surfaces. Let a plane wave be incident upon the slab in the $\hat{y}$ direction. We neglect reflections at the surfaces. Upon entering the medium, the wave is decomposed into an ordinary wave with $\overline{D} = \hat{x} D_o$ and an extraordinary wave with $\overline{D} = \hat{z} D_e$ along the optic axis. The spatial dependence of the $\overline{D}$ vector then becomes

$$\overline{D} = \hat{x} D_o e^{i\omega y/\sqrt{\nu\kappa}} + \hat{z} D_e e^{i\omega y/\sqrt{\nu\kappa_z}}$$

Notice that we used $\theta = \pi/2$ in determining $k = \omega/u$ by using (9) and (14). Thus after propagating a distance of y, the polarization of the incoming wave has been changed. If $D_o = D_e$, then at $y = 0$, the wave is linearly polarized in a direction 45° with respect to the optic

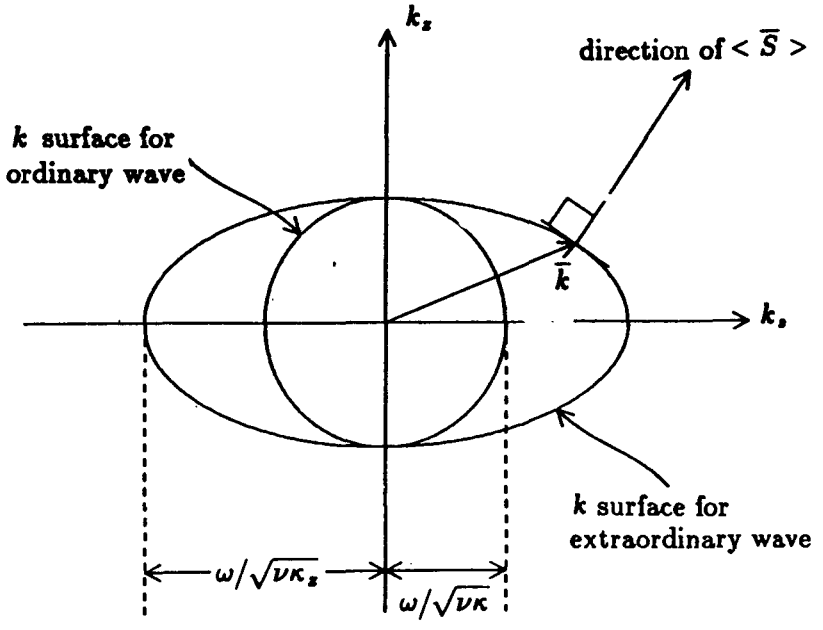


Figure 2.4.3 k **surfaces for the ordinary and extraordinary waves inside a positive uniaxial medium with** $\kappa/\kappa_z > 1$.

axis. After a distance d such that

$$\frac{\omega d}{\sqrt{\nu\kappa}} - \frac{\omega d}{\sqrt{\nu\kappa_z}} = \frac{(2m+1)\pi}{2}$$

where m is an integer, the wave becomes circularly polarized. A slab of such thickness is known as a quarter-wave plate. A polaroid is also a uniaxial medium where ϵ_z has a very large imaginary part such that the extraordinary wave is drastically attenuated after having been transmitted, whereas the ordinary wave component is transmitted with only a little attenuation.

b. k Surfaces

The dispersion relations in (9) and (14) for the ordinary and extraordinary waves can be converted from their angular dependence to the rectangular components of $\overline{k}$ in a three-dimensional k space with the axes formed by k_x, k_y and k_z. Notice that $u = \omega/k$, $k\cos\theta = k_z$, and $k\sin\theta = k_s$ where k_s is the transverse wavenumber representing the transverse component of $\overline{k}$.

We find that

$$\omega^2 = \nu\kappa k_x^2 + \nu\kappa k_z^2 \tag{19}$$

for the ordinary wave and

$$\omega^2 = \nu\kappa k_x^2 + \nu\kappa_z k_z^2 \tag{20}$$

for the extraordinary wave. Equation (19) describes a circle and (20) describes an ellipse [Fig. 2.4.3]. Rotating them about k_z, we obtain a sphere for the ordinary wave and an ellipsoid for the extraordinary wave.

The dispersion relations provide a functional relationship among components of the $\overline{k}$ vector and the angular frequency ω. The equations are usually quadratic in k for each characteristic wave when there are more than one. We write in general

$$f(k_x, k_y, k_z, \omega) = 0$$

In the three-dimensional $\overline{k}$ space, a quadratic equation describes a two-dimensional hypersurface. The surface is called the wave surface or simply the $\overline{k}$ surface.

We observe that the magnitude of $\overline{k}$, as described by the k surface, may vary as a function of its direction. In a particular direction, the $\overline{k}$ vector intersects the k surface at a point. The magnitude of $\overline{k}$ in this direction is proportional to the length from the origin to this point. The phase velocity of a wave in the media is $u = \omega/k$. The group velocity pertaining to the k surface is defined by (2.2.16) :

$$v_{gi} = \frac{\partial\omega}{\partial k_i} = -\frac{\partial f/\partial k_i}{\partial f/\partial\omega}$$

Since a vector with component $\partial f/\partial k_i$ is normal to the k surface $f(k_x, k_y, k_z, \omega) = 0$, we conclude that, as a wave propagates with phase velocity u along the direction of $\overline{k}$ that intersects the wave surface at a point, the group velocity is in a direction normal to the k surface at that point.

We now prove that in general the direction of flow of the time-average Poynting's vector is also normal to the tangent of the k surfaces. Thus the group velocity is in the direction of the energy velocity which is defined as the time-average Poynting's vector divided by the electromagnetic energy density. Mathematically we wish to show

that $<\overline{S}>$ is perpendicular to the tangent of the k surface, namely $\delta\overline{k} \cdot <\overline{S}> = 0$. First we take the differential of $\overline{k} \times \overline{E} = \omega\overline{B}$ and $\overline{k} \times \overline{H}^* = -\omega\overline{D}^*$ and obtain

$$\delta\overline{k} \times \overline{E} + \overline{k} \times \delta E = \omega\,\delta\overline{B} \tag{21}$$

$$\delta\overline{k} \times \overline{H}^* + \overline{k} \times \delta\overline{H}^* = -\omega\,\delta\overline{D}^* \tag{22}$$

We dot-multiply (21) by $\overline{H}^*$ and (22) by $\overline{E}$ and then subtract. Using the vector identity $\overline{A} \cdot (\overline{B} \times \overline{C}) = \overline{B} \cdot (\overline{C} \times \overline{A}) = \overline{C} \cdot (\overline{A} \times \overline{B})$ we obtain

$$\begin{aligned} 2\delta\overline{k} \cdot (\overline{E} \times \overline{H}^*) &= \omega(\overline{H}^* \cdot \delta\overline{B} + \overline{E} \cdot \delta\overline{D}^*) + \delta\overline{E} \cdot (\overline{k} \times \overline{H}^*) - \delta\overline{H}^* \cdot (\overline{k} \times \overline{E}) \\ &= \omega\left\{\overline{H}^* \cdot \delta\overline{B} + \overline{E} \cdot \delta\overline{D}^* - \overline{D}^* \cdot \delta\overline{E} - \overline{B} \cdot \delta\overline{H}^*\right\} \end{aligned} \tag{23}$$

For the uniaxial medium, we find

$$\begin{aligned} &2\,\delta\overline{k} \cdot (\overline{E} \times \overline{H}^*) \\ &\quad = \omega\left\{\overline{E} \cdot \overline{\overline{\epsilon}}^* \cdot \delta\overline{E}^* - \delta\overline{E} \cdot \overline{\overline{\epsilon}}^* \cdot \overline{E}^* + \overline{H}^* \cdot \mu\,\delta\overline{H} - \delta\overline{H}^* \cdot \mu\overline{H}\right\} \\ &\quad = \omega\left\{[\overline{E}^* \cdot \overline{\overline{\epsilon}} \cdot \delta\overline{E}]^* - [\overline{E}^* \cdot \overline{\overline{\epsilon}}^+ \cdot \delta\overline{E}] + [\overline{H}^* \cdot \mu\,\delta\overline{H}] - [\overline{H}^* \cdot \mu^*\delta\overline{H}]^*\right\} \end{aligned}$$

Using the lossless condition $\overline{\overline{\epsilon}}^+ = \overline{\overline{\epsilon}}$ and $\mu^* = \mu$, we observe that the right-hand side is a pure imaginary quantity. Since the time-average Poynting's vector is equal to one-half times the real part of $\overline{E} \times \overline{H}^*$, we conclude that

$$\delta\overline{k} \cdot <\overline{S}> = 0 \tag{24}$$

In general we can show that the right-hand side of (23) is a pure imaginary quantity for lossless bianisotropic media. Thus, at a point on the k surface, the direction of the time-average Poynting's vector is normal to the surface at that point.

2.5 Plane Waves in Gyrotropic Media

As another example of the application of the kDB system to the solution of characteristic waves inside homogeneous media, consider a gyrotropic medium possessing the following constitutive relations:

$$\overline{H} = \nu\overline{B} \tag{1}$$

$$\overline{E} = \overline{\overline{\kappa}} \cdot \overline{D} \tag{2}$$

where

$$\overline{\overline{\kappa}} = \begin{pmatrix} \kappa & i\kappa_g & 0 \\ -i\kappa_g & \kappa & 0 \\ 0 & 0 & \kappa_z \end{pmatrix} \tag{3}$$

One example of such a gyrotropic medium is an anisotropic plasma with an externally applied dc magnetic field in the $\hat{z}$ direction. The constitutive parameters κ, κ_g and κ_z have been derived and expressed in terms of the plasma frequency and cyclotron frequency as given in Problem P2.1.

a. Dispersion Relations and Characteristic Waves

We transform the constitutive matrices to the kDB system by applying (2.3.25)–(2.3.28). We find $\nu_k = \nu$ and

$$\begin{aligned} \overline{\overline{\kappa}}_k &= \overline{\overline{T}} \cdot \overline{\overline{\kappa}} \cdot \overline{\overline{T}}^{-1} \\ &= \begin{pmatrix} \kappa & i\kappa_g \cos\theta & i\kappa_g \sin\theta \\ -i\kappa_g \cos\theta & \kappa\cos^2\theta + \kappa_z \sin^2\theta & (\kappa - \kappa_z)\sin\theta\cos\theta \\ -i\kappa_g \sin\theta & (\kappa - \kappa_z)\sin\theta\cos\theta & \kappa\sin^2\theta + \kappa_z\cos^2\theta \end{pmatrix} \end{aligned} \tag{4}$$

Substituting the values for the constitutive elements in (2.3.36)-(2.3.37), we obtain

$$\begin{pmatrix} \kappa & i\kappa_g\cos\theta \\ -i\kappa_g\cos\theta & \kappa\cos^2\theta + \kappa_z\sin^2\theta \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = \begin{pmatrix} 0 & u \\ -u & 0 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \tag{5}$$

$$\nu \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} 0 & -u \\ u & 0 \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} \tag{6}$$

Eliminating $\overline{B}_k$ yields

$$\begin{pmatrix} u^2 - \nu\kappa & -i\nu\kappa_g\cos\theta \\ i\nu\kappa_g\cos\theta & u^2 - \nu(\kappa\cos^2\theta + \kappa_z\sin^2\theta) \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = 0 \tag{7}$$

For nontrival solutions for $\overline{D}_k$, we set the determinant of the 2×2 matrix equal to zero and obtain

$$u^2 = \frac{\nu}{2}\left[\kappa(1+\cos^2\theta) + \kappa_z\sin^2\theta \pm \sqrt{(\kappa-\kappa_z)^2\sin^4\theta + 4\kappa_g^2\cos^2\theta}\right] \tag{8}$$

In terms of the components of the $\overline{k}$ vector, we have

$$\omega^2 = \frac{\nu}{2}\left[\kappa(k^2 + k_z^2) + \kappa_z k_s^2 \pm \sqrt{(\kappa - \kappa_z)^2 k_s^4 + 4\kappa_g^2 k_z^2 k^2}\right]$$

This is the dispersion relation relating ω and $\overline{k}$.

The two components of the field vector $\overline{D}_k$ are related by

$$\frac{D_2}{D_1} = \frac{-2i\kappa_g \cos\theta}{(\kappa - \kappa_z)\sin^2\theta \pm \sqrt{(\kappa - \kappa_z)^2 \sin^4\theta + 4\kappa_g^2 \cos^2\theta}} \tag{9}$$

The expression can be greatly simplified if we define an angle ψ such that

$$\tan 2\psi = \frac{2\kappa_g \cos\theta}{(\kappa - \kappa_z)\sin^2\theta} \tag{10}$$

We find that for the characteristic wave with the phase velocity u having the plus sign in (8), (9) becomes

$$\frac{D_2}{D_1} = -i\tan\psi \tag{11}$$

We call it the Type I wave in the gyrotropic medium. For the characteristic wave with the phase velocity u having the minus sign in (8), we find from (9)

$$\frac{D_2}{D_1} = i\cot\psi \tag{12}$$

We call it the Type II wave. Both characteristic waves are elliptically polarized. Assume that $\kappa > \kappa_z$ and κ_g is positive, then the Type I wave is left-hand polarized and the Type II wave is right-hand polarized. When κ_g is zero, the medium becomes uniaxial and the characteristic waves become linearly polarized. Both characteristic waves are also linearly polarized when the wave propagation direction is perpendicular to $\hat{z}$ and thus $\theta = \pi/2$. This birefringence is known as the Cotton-Mouton effect.

b. Faraday Rotation

When the wave propagation direction is along $\hat{z}$, we have $\theta = 0$ and (7) becomes

$$\begin{pmatrix} u^2 - \nu\kappa & -i\nu\kappa_g \\ i\nu\kappa_g & u^2 - \nu\kappa \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = 0 \tag{13}$$

The magnitudes of the phase velocity are seen to be

$$u = \sqrt{\nu(\kappa \pm \kappa_g)} \tag{14}$$

and the ratio of the components of $\overline{D}_k$ is

$$\frac{D_2}{D_1} = \mp i \tag{15}$$

Thus, both characteristic waves are circularly polarized. The left-hand circularly polarized wave has a velocity $(\nu\kappa + \nu\kappa_g)^{1/2}$ and the right-hand, circularly polarized wave has a velocity $(\nu\kappa - \nu\kappa_g)^{1/2}$.

Consider a linearly polarized plane wave entering a gyrotropic medium along the $\hat{z}$ direction, neglecting reflections at the boundaries. The incoming wave is decomposed into two circularly polarized waves propagating at different velocities. We let

$$\overline{D} = \hat{e}_1 D_o = \frac{D_o}{2}(\hat{e}_1 + \hat{e}_2 i) + \frac{D_o}{2}(\hat{e}_1 - \hat{e}_2 i) \tag{16}$$

where D_o is a real number. After traveling a distance z_0 inside the medium, the two waves are phase-shifted by different amounts,

$$\begin{aligned}\overline{D} &= \frac{D_o}{2}(\hat{e}_1 + \hat{e}_2 i)e^{i\phi_r} + \frac{D_o}{2}(\hat{e}_1 - \hat{e}_2 i)e^{i\phi_l} \\ &= \hat{e}_1 \frac{D_o}{2}(e^{i\phi_r} + e^{i\phi_l}) + \hat{e}_2 \frac{iD_o}{2}(e^{i\phi_r} - e^{i\phi_l})\end{aligned} \tag{17}$$

where

$$\phi_r = \frac{\omega z_0}{\sqrt{\nu(\kappa - \kappa_g)}} \tag{18}$$

$$\phi_l = \frac{\omega z_0}{\sqrt{\nu(\kappa + \kappa_g)}} \tag{19}$$

For the ratio of the two components of $\overline{D}_k$, we find

$$\frac{D_2}{D_1} = i\frac{e^{i\phi_r} - e^{i\phi_l}}{e^{i\phi_r} + e^{i\phi_l}} = -\tan\frac{(\phi_r - \phi_l)}{2}$$

The two components are in phase and the wave is linearly polarized. Notice that the incoming wave is rotated clockwise by an angle $(\phi_r - \phi_l)/2$.

Now consider the case where the wave is propagating along the $-\hat{z}$ direction. We have $\theta = \pi$ and (7) becomes

$$\begin{pmatrix} u^2 - \nu\kappa & -i\nu\kappa_g \\ i\nu\kappa_g & u^2 - \nu\kappa \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = 0 \tag{20}$$

For phase velocities

$$u = \sqrt{\nu(k \pm k_g)} \tag{21}$$

we have

$$\frac{D_2}{D_1} = \pm i \tag{22}$$

The upper sign corresponds to left-hand circular polarization and the lower sign to right-hand circular polarization. For a linearly polarized wave as shown in (16) to propagate to a distance $z = -z_0$,

$$\overline{D} = \frac{D_o}{2}(\hat{e}_1 + \hat{e}_2 i)e^{-i\phi_l} + \frac{D_o}{2}(\hat{e}_1 - \hat{e}_2 i)e^{-i\phi_r} \tag{23}$$

We find

$$\frac{D_2}{D_1} = i\frac{e^{-i\phi_l} - e^{-i\phi_r}}{e^{-i\phi_l} + e^{-i\phi_r}} = -\tan\frac{\phi_r - \phi_l}{2} \tag{24}$$

Again we see that the incoming wave is rotated clockwise by an angle $(\phi_r - \phi_l)/2$. Thus, irrespective of whether the wave is propagating in the $+\hat{z}$ or $-\hat{z}$ direction, the wave is rotated in the same direction by the same angle.

The phenomenon of rotation of a linearly polarized field vector when passing through a gyrotropic medium is known as *Faraday rotation*. For a plasma medium, the electrons circulating along the magnetic field lines are responsible for this effect. Faraday rotation also occurs in ferrites in the presence of external magnetic fields; there the effect is caused by precession of spin axes around the magnetic field. A parallel analysis can be carried out for ferrites by using a magnetically anisotropic model with an impermeability tensor $\overline{\overline{\nu}}$.

2.6 Plane Waves in Bianisotropic Media

Consider bianisotropic media with the following constitutive relations:

$$\overline{E} = \begin{pmatrix} \kappa & 0 & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & \kappa_z \end{pmatrix} \cdot \overline{D} + \begin{pmatrix} \chi & 0 & 0 \\ 0 & \chi & 0 \\ 0 & 0 & \chi_z \end{pmatrix} \cdot \overline{B} \tag{1}$$

$$\overline{H} = \begin{pmatrix} \gamma & 0 & 0 \\ 0 & \gamma & 0 \\ 0 & 0 & \gamma_z \end{pmatrix} \cdot \overline{D} + \begin{pmatrix} \nu & 0 & 0 \\ 0 & \nu & 0 \\ 0 & 0 & \nu_z \end{pmatrix} \cdot \overline{B} \tag{2}$$

When $\overline{\overline{\chi}} = \overline{\overline{\gamma}}$, this relation reduces to that used by Dzyaloshinskii in his description of magnetoelectric media.

a. Dispersion Relations and Characteristic Waves

In the kDB system, the constitutive matrix $\overline{\overline{\kappa}}_k$ becomes

$$\overline{\overline{\kappa}}_k = \begin{pmatrix} \kappa & 0 & 0 \\ 0 & \kappa\cos^2\theta + \kappa_z \sin^2\theta & (\kappa - \kappa_z)\sin\theta\cos\theta \\ 0 & (\kappa - \kappa_z)\sin\theta\cos\theta & \kappa\sin^2\theta + \kappa_z\cos^2\theta \end{pmatrix} \tag{3}$$

A similar form holds for the other matrices, $\overline{\overline{\chi}}_k$, $\overline{\overline{\gamma}}_k$ and $\overline{\overline{\nu}}_k$. Inserting the corresponding constitutive parameters in (2.3.36) and (2.3.37) and eliminating $\overline{B}_k$, we obtain

$$\begin{pmatrix} \kappa_\theta(u^2\nu + \nu_\theta\chi\gamma - \kappa\nu\nu_\theta) & u\kappa_\theta(\nu_\theta\chi - \nu\gamma_\theta) \\ u\kappa(\nu_\theta\gamma - \nu\chi_\theta) & \kappa(u^2\nu_\theta + \nu\chi_\theta\gamma_\theta - \kappa_\theta\nu\nu_\theta) \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = 0 \tag{4}$$

where we use this short notation:

$$\begin{aligned} \kappa_\theta &= \kappa\cos^2\theta + \kappa_z\sin^2\theta \\ \nu_\theta &= \nu\cos^2\theta + \nu_z\sin^2\theta \\ \chi_\theta &= \chi\cos^2\theta + \chi_z\sin^2\theta \\ \gamma_\theta &= \gamma\cos^2\theta + \gamma_z\sin^2\theta \end{aligned}$$

Solving for u and $\overline{D}_k$ from (4), although tedious, is straightforward. We shall now discuss several special cases.

Consider a lossless magnetoelectric medium in which $\overline{\overline{\gamma}} = \overline{\overline{\chi}}$ are both real. We see from (4) that the characteristic waves are linearly polarized. In the $\hat{z}$ direction $\theta = 0$, and (4) becomes

$$\begin{pmatrix} u^2 - \kappa\nu + \chi^2 & 0 \\ 0 & u^2 - \kappa\nu + \chi^2 \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = 0 \tag{5}$$

This is a degenerate case. The characteristic waves can have any polarization, and the phase velocity is $u^2 = \kappa\nu - \chi^2$. Note that we must have $\kappa\nu > \chi^2$; otherwise the velocity becomes imaginary.

b. Optical Activity

Consider the case with both χ and γ imaginary and the bianisotropic medium lossless. Let $\overline{\overline{\chi}} \to i\overline{\overline{\chi}}$; then the lossless condition requires $\overline{\overline{\gamma}} = -i\overline{\overline{\chi}}$. We see from (4) that the characteristic waves are elliptically polarized. In the $\hat{z}$ direction (4) becomes

$$\begin{bmatrix} u^2 - \kappa\nu + \chi^2 & i2\chi u \\ -i2\chi u & u^2 - \kappa\nu + \chi^2 \end{bmatrix} \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} = 0 \tag{6}$$

The velocity of the propagation is determined from $u^2 - \kappa\nu + \chi^2 = \pm 2\chi u$, and the corresponding polarization is $D_2/D_1 = \pm i$. Thus both characteristic waves are circularly polarized. The right-hand circularly polarized wave has a velocity $\sqrt{\kappa\nu} + \chi$, and the left-hand circularly polarized wave has a velocity $\sqrt{\kappa\nu} - \chi$. As in the case of gyrotropic media, a linearly polarized wave entering this medium along the $\hat{z}$ direction is broken up into two characteristic waves that propagate at different velocities. The net result is a rotation of the propagation vector. A profound difference exists, however, between this rotation and the Faraday rotation. A comparison of (4) with (2.5.7) reveals that the off-diagonal elements in (2.5.7) change sign when we change θ from 0 to π, while those in (4) remain unchanged.

The significance of this difference can be demonstrated as follows: consider a linearly polarized wave that propagates through a slab of gyrotropic medium along the $\hat{z}$ direction. Assume that, upon exiting, its polarization vector is rotated 45°. If the wave is reflected by a mirror and re-enters the slab, the polarization vector is rotated a total of 90° after the whole journey. Consider the same experiment with the gyrotropic medium replaced by a bianisotropic medium as discussed above. On its return path after being reflected by the mirror, the polarization vector is rotated back to its original position and the net result is no rotation at all. Because of this difference, we call this rotatory power *optical activity* to distinguish it from the Faraday effect. As we shall see later, in Chapter V, the optical activity is reciprocal, whereas the Faraday effect is nonreciprocal.

2.7 Plane Waves in Nonlinear Media

Consider a nonlinear medium characterized by the constitutive relation

$$\overline{D}(\overline{r}, t) = \epsilon_o \overline{E}(\overline{r}, t) + \overline{P}(\overline{r}, t) \tag{1a}$$

with the ith component of $\overline{P}$

$$P_i = \chi_{ij} E_j + 2\chi_{ijk} E_j E_k + 4\chi_{ijkl} E_j E_k E_l + \cdots \tag{1b}$$

We have studied in previous sections the linear term χ_{ij}. The second-order nonlinear term χ_{ijk} is responsible for the phenomena of second-harmonic generation, and parametric amplification and oscillation. The third-order nonlinear term χ_{ijkl} gives rise to the effects of third-harmonic generation, Raman and Brillouin scattering, self-focusing, and phase conjugation.

The space–time-dependent Maxwell's equations in source-free regions read

$$\nabla \times \overline{E}(\bar{r},t) = -\frac{\partial}{\partial t}\overline{B}(\bar{r},t) \tag{2a}$$

$$\nabla \times \overline{H}(\bar{r},t) = -\frac{\partial}{\partial t}\overline{D}(\bar{r},t) \tag{2b}$$

Specializing to one dimension by letting $\partial/\partial x = \partial/\partial y = 0$, we obtain the wave equation for the ith component of $\overline{E}$

$$\frac{\partial^2}{\partial z^2} E_i - \mu_o \epsilon_o \frac{\partial^2}{\partial t^2} E_i - \mu_o \frac{\partial^2}{\partial t^2} P_i = 0 \tag{3}$$

We assume plane wave solutions of frequency dependence ω_1, ω_2, and ω_3.

$$E_{1i}(z,t) = \frac{1}{2}\left\{E_{1i}(z)e^{i(k_1 z - \omega_1 t)} + c.c.\right\} \tag{4a}$$

$$E_{2j}(z,t) = \frac{1}{2}\left\{E_{2j}(z)e^{i(k_2 z - \omega_2 t)} + c.c.\right\} \tag{4b}$$

$$E_{3k}(z,t) = \frac{1}{2}\left\{E_{3k}(z)e^{i(k_3 z - \omega_3 t)} + c.c.\right\} \tag{4c}$$

where $k_l = \omega_l(\mu_o \epsilon_l)^{1/2}$ with $l = 1,2,3$ and $c.c.$ denotes complex conjugate. When $\overline{P} = 0$, the amplitudes $E_{1i}(z), E_{2j}(z)$, and $E_{3k}(z)$ will be independent of z.

a. Second-Harmonic Generation (SHG)

Consider only the second-order nonlinear term χ_{ijk}. For the electric fields at ω_3 and ω_2, we have

$$P_i(z,t) = 2\chi_{ijk}\left\{\frac{1}{2}E_{3j}e^{i(k_3 z-\omega_3 t)} + \frac{1}{2}E_{2j}e^{i(k_2 z-\omega_2 t)} + c.c.\right\} \cdot \left\{\frac{1}{2}E_{3k}e^{i(k_3 z-\omega_3 t)} + \frac{1}{2}E_{2k}e^{i(k_2 z-\omega_2 t)} + c.c.\right\} \quad (5)$$

For $\omega_1 = \omega_3 - \omega_2$, the nonlinear polarization is

$$P_{1i}(z,t) = \frac{1}{2}\chi_{ijk}\left\{(E_{3j}E_{2k}^* + E_{3k}E_{2j}^*)e^{i(k_3-k_2)z-i(\omega_3-\omega_2)t} + c.c.\right\} = \chi_{ijk}\left\{E_{3j}E_{2k}^* e^{i(k_3-k_2)z-i(\omega_3-\omega_2)t} + c.c.\right\} \quad (6)$$

where use is made of the summation convention and the lossless condition of $\chi_{ijk} = \chi_{ikj}$.

We assume small variation of E_{1i} as a function of z such that $d^2E_{1i}/dz^2 \ll k_1 dE_{1i}/dz$. Letting $k_l = \omega_l(\mu_o\epsilon_l)^{1/2}$ and $\epsilon_{ij} = \epsilon_o + \chi_{ij} = \epsilon_l\delta_{ij}$ for E_{li} components, we reduce the wave equation in (3) for $E_{1i}(z,t)$ to the following for the complex $E_{1i}(z)$

$$\frac{d}{dz}E_{1i}(z) - i\omega_1\sqrt{\frac{\mu_o}{\epsilon_1}}\chi_{ijk}E_{3j}(z)E_{2k}^*(z)e^{i(-k_1+k_3-k_2)z} = 0 \quad (7a)$$

Similarly, we find

$$\frac{d}{dz}E_{3j}(z) - i\omega_3\sqrt{\frac{\mu_o}{\epsilon_3}}\chi_{jkl}E_{1k}(z)E_{2l}(z)e^{i(k_1-k_3+k_2)z} = 0 \quad (7b)$$

$$\frac{d}{dz}E_{2k}^*(z) + i\omega_2\sqrt{\frac{\mu_o}{\epsilon_2}}\chi_{klm}E_{1l}(z)E_{3m}^*(z)e^{i(k_1-k_3+k_2)z} = 0 \quad (7c)$$

For second-harmonic generation (SHG), $\omega_1 = \omega_2$ and $\omega_3 = \omega_1 + \omega_2 = 2\omega_1$. Equation (7*c*) is merely the complex conjugate of (7*a*). Equation (7*b*), however, should be rederived from (5) which now takes the form

$$P_i(z,t) = 2\chi_{ijk}\left\{\frac{1}{2}E_{1j}\,e^{i(k_1 z-\omega_1 t)} + c.c.\right\}\left\{\frac{1}{2}E_{1k}\,e^{i(k_1 z-\omega_1 t)} + c.c.\right\}$$

which yield for $\omega_3 = 2\omega_1$,

$$P_{3i}(z,t) = \frac{1}{2}\chi_{ijk}\left\{E_{1j}E_{1k}\,e^{i2(k_1 z-\omega_1 t)} + c.c.\right\}$$

Comparison with (6) shows a factor of $1/2$ difference. The wave equation (3) that led to $(7b)$ now becomes

$$\frac{d}{dz}E_{3j}(z) - i\frac{\omega_3}{2}\sqrt{\frac{\mu_o}{\epsilon_3}}\chi_{ikl}E_{1k}(z)E_{1l}(z)e^{-i\Delta kz} = 0 \qquad (8a)$$

Equation $(7a)$ can be rewritten as

$$\frac{d}{dz}E_{1i}(z) - i\omega_1\sqrt{\frac{\mu_o}{\epsilon_1}}\chi_{ijk}E_{3j}(z)E^*_{1k}(z)e^{i\Delta kz} = 0 \qquad (8b)$$

where $\Delta k = k_3 - 2k_1$. Equation (8) forms the basis for the study of SHG. Notice that $k_3 = \omega_3(\mu_o\epsilon_3)^{1/2}$ and $k_1 = \omega_1(\mu_o\epsilon_1)^{1/2}$, E_{3j} is the jth component of the electric field at frequency ω_3, and E_{1i} is the ith component of the electric field at frequency ω_1.

Assuming weak second-harmonic generation such that the depletion of wave at ω_1 is small. In $(8b)$ $E_{3j} \approx 0$ and the solution for E_{1i} is a constant. Let there be zero second-harmonic input at $z = 0$ such that $E_{3j}(0) = 0$, we find from $(8a)$ the approximate solution

$$E_{3j}(z) = \frac{\omega_3}{2}\sqrt{\frac{\mu_o}{\epsilon_3}}\chi_{jkl}E_{1k}E_{1l}\frac{1 - e^{-i\Delta kz}}{\Delta k}$$

It is seen that the power generated at ω_3 contains the interference factor $\sin^2(\Delta kz/2)$. Thus the region of z for generation of second-harmonic wave should be smaller than the coherence length defined by $l_c = 2\pi/\Delta k$.

The coherence length is infinite and the second-harmonic generation is most effective when $\Delta k = 0$, which is known as the phase-matching condition. Notice that the electric field E_{3j} at ω_3 can be polarized differently from E_{1k} at ω_1. It suggests that the phase-matching condition can be met by using dispersive anisotropic media. For instance E_{1k} can be an extraordinary wave at ω_1 and E_{3j} an ordinary wave at ω_3. The phase-matching condition $\Delta k = \omega_3\sqrt{\mu_o\epsilon_3(\omega_3)} - 2\omega_1\sqrt{\mu_o\epsilon_1(\omega_1)} = 0$ with $\omega_3 = 2\omega_1$ gives $n_o(\omega_3) - n_e(\omega_1) = 0$ along

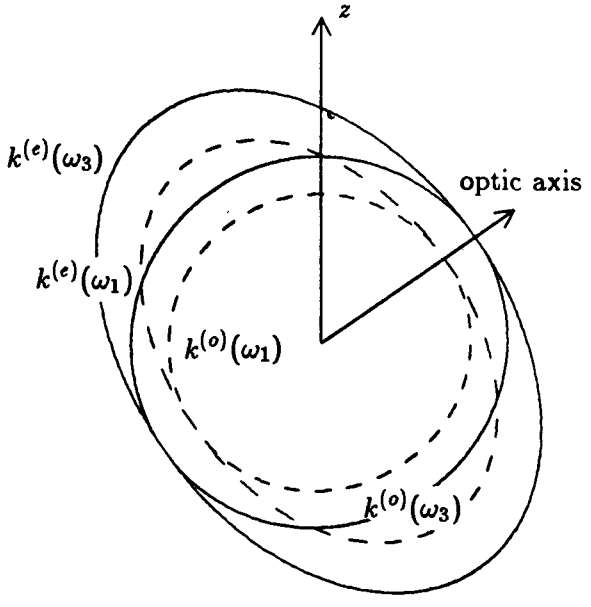


Figure 2.7.1 Phase matching.

the $\hat{z}$ direction, where $n_o(\omega_3) = c\sqrt{\mu\epsilon_3(\omega_3)}$ and $n_e(\omega_1) = c\sqrt{\mu\epsilon_1(\omega_1)}$. Let $k^{(o)}(\omega_3) = \omega\sqrt{\mu\epsilon_3(\omega_3)}$ and $k^{(e)}(\omega_1) = \omega\sqrt{\mu\epsilon_1(\omega_1)}$. We show in Figure 2.7.1 that the k surfaces at ω_1 and ω_3 can be oriented to intersect on the z axis to satisfy the phase-matching condition.

When the phase-matching condition of $\Delta k = 0$ is satisfied, (8) can be simplified by letting $E_{3j} = i\sqrt{\omega_3/n_o(\omega_3)}A_3$, $E_{1k} = \sqrt{\omega_1/n_e(\omega_1)}A_1$ and $\kappa = c\mu\chi_{211}\omega_1\sqrt{\omega_3}/n_e\sqrt{n_o}$. We obtain, assuming A_1 to be real,

$$\frac{d}{dz}A_3 = \frac{\kappa}{2}A_1^2 \tag{9a}$$

$$\frac{d}{dz}A_1 = -\kappa A_1 A_3 \tag{9b}$$

Summing (9a) multiplied by $2A_3$ and (9b) multiplied by A_1 yields

$$\frac{d}{dz}(A_1^2(z) + 2A_3^2(z)) = 0$$

Since there is no input at ω_3, $A_3(z = 0) = 0$, and we find $A_1^2(z) + 2A_3^2(z) = A_1^2(0)$. Equation (9$a$) becomes

$$\frac{d}{dz}A_3(z) = \frac{\kappa}{2}\left[A_1^2(0) - 2A_3^2(z)\right]$$

and the solution is

$$A_3(z) = \frac{A_1(0)}{\sqrt{2}} \tanh\left[\frac{A_1(0)}{\sqrt{2}}\kappa z\right] \tag{10}$$

Notice that as $z \to \infty$, $A_3 \to A_1(0)/\sqrt{2}$ and the power generated at the second-harmonic of $\omega_3 = 2\omega_1$ is equal to half of the input power at ω_1 after total conversion is complete. This satisfies energy conservation as the photon energy at ω_3 is $\hbar\omega_3 = 2\hbar\omega_1$.

b. Phase Conjugation

We now consider phase conjugation caused by four-wave mixing as a result of the third-order nonlinear term χ_{ijkl}. The objective is to generate an output wave at ω_3 which is the phase conjugate of an input wave at ω_4. A nonlinear medium is pumped by waves at ω_1 and ω_2 in opposite directions with $\bar{k}_1 = -\bar{k}_2$ [Fig. 2.7.2]. Assuming plane wave solutions, we find similar to (5) a nonlinear polarization due to the electric fields at ω_1, ω_2 and ω_4,

$$\begin{aligned} P_i(z,t) = \frac{1}{2}\chi_{ijkl} & \\ \cdot \left\{E_{1j}e^{i(\bar{k}_1\cdot\bar{r}-\omega_1 t)} + E_{2j}e^{i(\bar{k}_2\cdot\bar{r}-\omega_2 t)} + E_{3j}e^{i(\bar{k}_4\cdot\bar{r}-\omega_4 t)} + c.c.\right\} & \\ \cdot \left\{E_{1k}e^{i(\bar{k}_1\cdot\bar{r}-\omega_1 t)} + E_{2k}e^{i(\bar{k}_2\cdot\bar{r}-\omega_2 t)} + E_{3k}e^{i(\bar{k}_4\cdot\bar{r}-\omega_4 t)} + c.c.\right\} & \\ \cdot \left\{E_{1l}e^{i(\bar{k}_1\cdot\bar{r}-\omega_1 t)} + E_{2l}e^{i(\bar{k}_2\cdot\bar{r}-\omega_2 t)} + E_{3l}e^{i(\bar{k}_4\cdot\bar{r}-\omega_4 t)} + c.c.\right\} & \end{aligned} \tag{11}$$

Letting $\omega_1 = \omega_2 = \omega_4 = \omega$, $\bar{k}_1 = -\bar{k}_2$, and $\bar{k}_4 = -\hat{z}k$, we obtain for $\omega_3 = \omega_1 + \omega_2 - \omega_4$,

$$\begin{aligned} P_{3i}(z,t) &= \frac{1}{2}\chi_{ijkl}E_{1j}E_{2k}E_{4l}^* e^{i[(\bar{k}_1+\bar{k}_2)\cdot\bar{r}+kz-(\omega_1+\omega_2)t+\omega_4 t]} + c.c. \\ &= \frac{1}{2}\chi E_1 E_2 E_4^* e^{i(kz-\omega t)} + c.c. \end{aligned} \tag{12}$$

where we ignore the subscripts and set $\chi = 6\chi_{ijkl}$. The nonlinear polarization generates a wave $[\frac{1}{2}E_3(z)e^{i(kz-\omega t)} + c.c.]$ according to the wave equation (3) which now gives, ignoring the $\partial^2 E_3/\partial z^2$ term

$$ik\frac{d}{dz}E_3(z) = -\frac{1}{2}\omega^2\mu_o\chi E_1 E_2 E_4^* \tag{13}$$

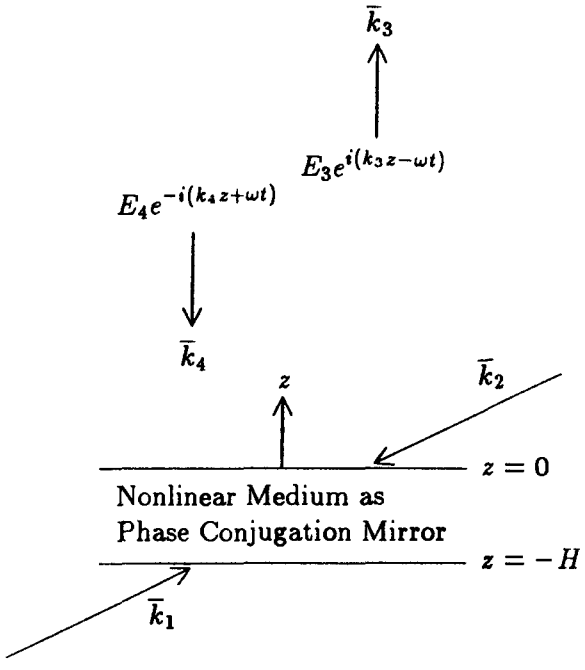


Figure 2.7.2 Generation of phase-conjugated waves.

where $k=\omega(\mu_o\epsilon)^{1/2}$ with $\epsilon_{ij}=\epsilon_o+\chi_{ij}=\epsilon\delta_{ij}$ and the second-order term $\chi_{ijk}=0$.

The newly created E_3 wave will mix with E_1 and E_2 to generate the polarization

$$P_4(z,t)=\frac{1}{2}\chi E_1 E_2 E_3^* e^{-i(kz+\omega t)}+c.c. \tag{14}$$

which interacts strongly with $E_4 e^{-i(kz+\omega t)}$. The wave equation for $E_4(z)$ gives rise to

$$-ik\frac{d}{dz}E_4(z)=-\frac{1}{2}\omega^2\mu_o\chi E_1 E_2 E_3^* \tag{15}$$

Equations (14)–(15) are obtained by a process similar to (11)–(13).

Defining a coupling coefficient

$$\kappa=\frac{\omega}{2}\sqrt{\frac{\mu_o}{\epsilon}}\chi E_1 E_2$$

and noticing that $k = \omega(\mu_o\epsilon)^{1/2}$, we obtain from (13) and the complex conjugate of (15) the coupled equations

$$\frac{d}{dz}E_3(z) = i\kappa E_4^*(z) \tag{16a}$$

$$\frac{d}{dz}E_4^*(z) = i\kappa^* E_3(z) \tag{16b}$$

Given $E_4(0)$ and $E_3(-H)$, the solutions to (16) take the form

$$E_3(z) = \frac{\cos|\kappa|z}{\cos|\kappa|H}E_3(-H) - i\frac{\kappa\sin|\kappa|(z+H)}{|\kappa|\cos|\kappa|H}E_4^*(0) \tag{17a}$$

$$E_4^*(z) = i\frac{|\kappa|\sin|\kappa|z}{\kappa\cos|\kappa|H}E_3(-H) - \frac{\cos|\kappa|(z+H)}{\cos|\kappa|H}E_4^*(0) \tag{17b}$$

Let $E_3(-H) = 0$. The reflected wave $E_3(0)$ at $z = 0$ due to an input wave $E_4(0)$ is

$$E_3(0) = -i\left(\frac{\kappa}{|\kappa|}\tan|\kappa|H\right)E_4^*(0) \tag{18}$$

Thus $E_3(0)$ is proportional to $E_4^*(0)$, the complex conjugate of $E_4(0)$.

Notice that the reflected wave is the conjugate of the incident wave in space but not in time. An incident wave with the pulse form $f(z + ct)$ will be of the form $f(z - ct)$ upon reflection from a conjugation mirror. Thus the reflected wave is the time reversal of the incident pulse. Furthermore if the input $E_3(0)$ is not a plane wave but has a complicated wavefront

$$E_4 = Re\left\{E_4(\bar{r})e^{-i(\omega t+kz)}\right\}$$

it follows that

$$E_3 = Re\left\{-i\left(\frac{\kappa}{|\kappa|}\tan|\kappa|H\right)E_4^*(\bar{r})e^{i(kz-\omega t)}\right\}$$

which is easily verified by modifying the above derivations on account of the linearity of the wave equations.

It is interesting to note that $|E_4(-H)| > |E_4(0)|$ and that for $\pi/4 < |\kappa|H < 3\pi/4$, $|E_3(0)| > |E_4(0)|$. The amplification of $E_4(-H)$ and the generation of $E_3(0)$ must be at the expense of the pump

waves E_1 and E_2. In fact, as $|\kappa|H = \pi/2$, $|E_4(-H)/E_4(0)| \to \infty$ and $|E_3(0)/E_4(0)| \to \infty$ which gives rise to natural resonance without an input wave.

PROBLEMS

Problem P2.1

Consider an electron plasma placed in a dc magnetic field $\overline{B}_0$.

(a) Show that the equation of motion for an electron is

$$-i\omega m\overline{v} = q(\overline{E} + \overline{v} \times \overline{B}_0)$$

under a time-harmonic excitation with angular frequency ω.

(b) Define a vector $\overline{\omega}_c = q\overline{B}_0/m$. Cross-multiply and dot-multiply the above equation by $\overline{B}_0$ to eliminate $\overline{\omega}_c \times \overline{v}$ and $\overline{\omega}_c \cdot \overline{v}$ and show that

$$N q\overline{v} = -i\omega\epsilon_o \cdot \left[\frac{-\omega_p^2}{\omega^2 - \omega_c^2}\overline{E} + \frac{\omega_p^2}{\omega^2(\omega^2 - \omega_c^2)}\overline{\omega}_c\overline{\omega}_c \cdot \overline{E} + i\frac{\omega_p^2}{\omega(\omega^2 - \omega_c^2)}\overline{\omega}_c \times \overline{E}\right]$$

Under the assumption that $\overline{B}_0$ is in the direction of the z axis, show that the permittivity tensor takes the following form

$$\overline{\overline{\epsilon}} = \begin{pmatrix} \epsilon & i\epsilon_g & 0 \\ -i\epsilon_g & \epsilon & 0 \\ 0 & 0 & \epsilon_z \end{pmatrix}$$

with elements

$$\epsilon = \epsilon_o\left[1 - \frac{\omega_p^2/\omega^2}{1 - \omega_c^2/\omega^2}\right]$$

$$\epsilon_g = \epsilon_o\left[\frac{-\omega_c\,\omega_p^2/\omega^3}{1 - \omega_c^2/\omega^2}\right]$$

$$\epsilon_z = \epsilon_o\left[1 - \frac{\omega_p^2}{\omega^2}\right]$$

Notice that $\omega_c = qB_0/m$ in this case is negative because the electron charge is negative.

(c) Show that for $\overline{E} = \overline{\overline{\kappa}} \cdot \overline{D}$

$$\overline{\overline{\kappa}} = \overline{\overline{\epsilon}}^{-1} = \begin{pmatrix} \kappa & i\kappa_g & 0 \\ -i\kappa_g & \kappa & 0 \\ 0 & 0 & \kappa_z \end{pmatrix}$$

where

$$\kappa = \frac{\epsilon}{\epsilon^2 - \epsilon_g^2} = \frac{1}{\epsilon_o}\left[\frac{1 - \omega_p^2/\omega^2 - \omega_c^2/\omega^2}{(1 - \omega_p^2/\omega^2)^2 - \omega_c^2/\omega^2}\right]$$

$$\kappa_g = \frac{-\epsilon_g}{\epsilon^2 - \epsilon_g^2} = \frac{1}{\epsilon_o}\left[\frac{\omega_c\omega_p^2/\omega^3}{(1 - \omega_p^2/\omega^2)^2 - \omega_c^2/\omega^2}\right]$$

$$\kappa_z = \frac{1}{\epsilon_z} = \frac{1}{\epsilon_o}\left[\frac{1}{1 - \omega_p^2/\omega^2}\right]$$

Problem P2.2

Consider an electron plasma with collisions, and introduce a collision frequency $\omega_{eff} \approx NT^{-3/2}$ such that a damping term is introduced and the force on the electron in the $\hat{x}$ direction becomes $f_x = d^2mx/dt^2 + \omega_{eff}\, d\,mx/dt$. Derive the constitutive relation and show that

$$\epsilon = \epsilon_o\left[1 - \frac{\omega_p^2}{\omega^2 + \omega_{eff}^2} + i\frac{\omega_p^2\omega_{eff}}{\omega(\omega^2 + \omega_{eff}^2)}\right]$$

Discuss the limiting case as $\omega_{eff}/\omega \to \infty$.

Problem P2.3

In this problem we examine dispersion in the vicinity of a resonant frequency. Show that, if an electron is originally bound to an ion as in an atom,

$$\left(\frac{d^2}{dt^2} + g\omega_0\frac{d}{dt} + \omega_0^2\right)\overline{P} = \frac{Ne^2}{m}\overline{E}$$

where $\overline{P} = Nq\overline{r}$ is the total dipole moment per unit volume, g is a damping constant, and ω_0 is the characteristic frequency of the

electron and accounts for the restoring force. Derive the complex permittivity

$$\epsilon(\omega) = \epsilon_o \left(1 + \frac{\omega_p^2}{\omega_0^2 - ig\omega\omega_0 - \omega^2}\right) = \epsilon_R(\omega) + i\epsilon_I(\omega)$$

and plot the real and imaginary parts of $\epsilon(\omega)$. Identify the region of normal dispersion, where ϵ_R increases with frequency, and the region of anomalous dispersion, where ϵ_R decreases with frequency. Show that ϵ_I is highest at the resonance frequency ω_0. Note: use the approximations $\omega \simeq \omega_0$, and $(\omega^2 - \omega_0^2) \simeq 2\omega_0(\omega - \omega_0)$.

Problem P2.4

In the macroscopic theory of dielectric dispersion, consider the equilibrium dipolar polarization to be represented by P_s. When an electric field E is applied, a distortion polarization P_1 is immediately established, but the remaining dipolar part of the polarization P_2 takes time to reach its equilibrium state. Letting the macroscopic relaxation time be τ, we have

$$\frac{dP_2}{dt} = \frac{1}{\tau}(P_s - P_1 - P_2)$$

where

$$P_s = (\epsilon_s - \epsilon_o)E$$
$$P_1 = (\epsilon_\infty - \epsilon_o)E$$

where ϵ_s and ϵ_∞ are both real and representative of the permittivity at static and infinite frequencies respectively.

(a) For a time-harmonic field with angular frequency ω, show that

$$P_2 = \frac{P_s - P_1}{1 - i\omega\tau}$$

and the macroscopic permittivity ϵ is

$$D = \epsilon E = \epsilon_o E + P_1 + P_2 = \left(\epsilon_\infty + \frac{\epsilon_s - \epsilon_\infty}{1 - i\omega\tau}\right) E$$

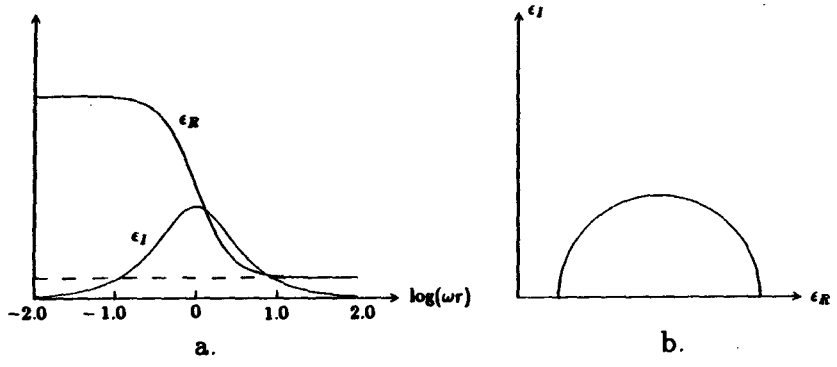


Figure P2.4

This is known as Debye's formula.

(b) Let $\epsilon = \epsilon_R + i\epsilon_I$ and plot ϵ_R and ϵ_I [Fig. P2.4a]. Label the values for ϵ_R and ϵ_I including the maximum point for ϵ_I. What are the numerical values for water molecules?

(c) Show that ϵ_I plotted as a function of ϵ_R is a circle. Find the radius of the Debye semicircle [Fig. P2.4b] and the points of intersection with the ϵ_R axis.

Problem P2.5

From Poynting's theorem in real space–time, identify

$$\overline{E}(\bar{r},t)\cdot\frac{\partial}{\partial t}\overline{D}(\bar{r},t)=\frac{\partial}{\partial t}W_e(\bar{r},t)+P_{EL}(\bar{r},t)$$

$$\overline{H}(\bar{r},t)\cdot\frac{\partial}{\partial t}\overline{B}(\bar{r},t)=\frac{\partial}{\partial t}W_m(\bar{r},t)+P_{ML}(\bar{r},t)$$

where

$W_e(\bar{r},t)$ = density of electric field energy

$W_m(\bar{r},t)$ = density of magnetic field energy

$P_{EL}(\bar{r},t)$ = density of electric polarization loss

$P_{ML}(\bar{r},t)$ = density of magnetic polarization loss

Considering narrow-band signals, use the Fourier transform to form the following equation:

$$\overline{E}(\bar{r},t)\cdot\frac{\partial}{\partial t}\overline{D}(\bar{r},t)=\frac{1}{2}\int d\omega\int d\omega'\,\overline{E}(\bar{r},\omega)\overline{E}^*(\bar{r},\omega')\cdot[i\omega'\epsilon^*(\omega')-i\omega\epsilon(\omega)]e^{-i(\omega-\omega')t}$$

(a) Show that for an isotropic dispersive media with real permittivity $\epsilon(\omega)$ and permeability μ, the time-average stored electric energy density is

$$< W_e >= \frac{1}{4}\frac{\partial(\omega\epsilon)}{\partial\omega}\left|\overline{E}\right|^2$$

and the time-average stored magnetic energy density is

$$< W_m >= \frac{1}{4}\mu\left|\overline{H}\right|^2$$

(b) Calculate the stored electric energy density for a plasma medium with permittivity $\epsilon = \epsilon_o[1-(\omega_p^2/\omega^2)]$ where $\omega_p^2 = Ne^2/\epsilon_o m$. Show that

$$< W_e >= \frac{\epsilon_o}{4}\left(1+\frac{\omega_p^2}{\omega^2}\right)\left|\overline{E}\right|^2$$

Identify the first term as the stored electric energy density in free space. Show that the second term represents the time-average kinetic energy density of the electric charges by noticing that $m\,d^2x/dt^2 = eE$.

Problem P2.6

Consider two real time-harmonic vectors $\overline{E}_1(t)$ and $\overline{E}_2(t)$ represented by complex spatial vectors $\overline{E}_1$ and $\overline{E}_2$ obeying the rule $\overline{E}(t) = Re\left\{\overline{E}e^{-i\omega t}\right\}$. Let $\overline{E}_1 = \hat{x} + \hat{y}\,i$ and $\overline{E}_2 = i(\hat{x} + \hat{y}\,i)$. Show that

(a) $\overline{E}_1 \times \overline{E}_2 = 0$ but $\overline{E}_1(t) \times \overline{E}_2(t) \neq 0$

(b) $\overline{E}_1 \cdot \overline{E}_2 = 0$ and $\overline{E}_1(t) \cdot \overline{E}_2(t) = 0$.

Problem P2.7

(a) The complex permittivity for bottom round steak is about $\epsilon = 40(1+i0.3)\epsilon_o$ at the operating frequency (2.5 GHz) of a microwave oven. What is the penetration depth?

(b) Calculate loss tangents and skin depths for sea water at frequencies 60 Hz and 10 MHz. Sea water can be characterized by conductivity $\sigma = 4\,\text{mho/m}$, permittivity $\epsilon = 80\epsilon_o$, and permeability $\mu = \mu_o$ at those frequencies.

(c) A 100-Hz electromagnetic wave is propagating down into the sea water with an electric field intensity E of 1 V/m just beneath the

sea surface. What is the intensity of E at a depth of 100 m? What are the time-average Poynting's power densities just beneath the surface and at a depth of 100 m?

Problem P2.8

Superconductivity was first observed by Kamerlingh Onnes in 1911. In 1933 Meissner and Ochsenfeld discovered that superconducting metals cannot be penetrated by magnetic fields. Magnetic fields are expelled from a normal metal when it is cooled to the superconducting state. The macroscopic theory of superconductivity was developed by London and London in 1935 followed by the microscopic theory of Bardeen, Cooper and Schrieffer in 1957.

A simple model of superconductivity calls for an electron plasma with a very high electron density N.

(a) Show that the penetration depth of a plasma with very large N takes the form

$$d_p = \sqrt{\frac{m}{Ne^2\mu_o}}$$

(b) Letting $N = 7 \times 10^{28}\,\mathrm{m}^{-3}$, calculate d_p.

(c) Compare the above result with the skin depth of a good conductor. Explain why a very slowly varying magnetic field can penetrate a good conductor but not a superconductor.

Problem P2.9

A sketch of a simple interferometer is shown in Figure P2.9. Light from a source is split into two beams by a semi-transparent mirror. The semi-transparent mirror has the property of transmitting $e^{-i\phi_t}/\sqrt{2}$ of the incident electric field and reflecting $e^{-i\phi_r}/\sqrt{2}$ of the incident field.

(a) Show that the semi-transparent mirror transmits and reflects 1/2 of the incident power.

(b) Consider the light source to be emitting light with field amplitude E_0 and assume $\epsilon = \epsilon_o$. Show by power conservation argument that

$$\phi_t - \phi_r = \frac{\pi}{2} + l\pi$$

(c) Suppose the region l_1 is filled with a plasma with electron den-

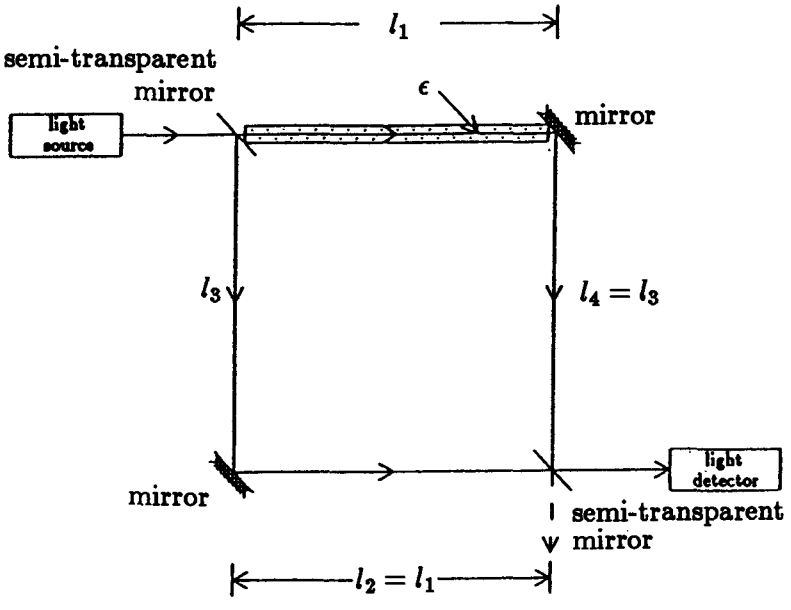


Figure P2.9

sity N. Remembering that

$$\epsilon = \epsilon_o \left(1 - \frac{\omega_p^2}{\omega^2}\right)$$

where $\omega_p = \sqrt{Ne^2/m\epsilon_o}$, find the power intensity as a function of N when $\omega_p \ll \omega$. Neglect reflections from the light detector and the plasma boundaries.

Problem P2.10

Pulsars, or pulsating radio stars, emit sharp bursts of radio energy about 5 to 50 milliseconds wide at intervals of about one second. For any given pulsar the repetition frequency is stable to within about one part in 10^8. The amplitude and shape of the pulses vary widely, but each pulsar has its own characteristic mean pulse profile.

Within a few months after the discovery of pulsars, distance estimates were obtained in the following manner. It was observed that the arrival time depends on the frequency of observation, the arrival time being later at lower frequencies. This delay is attributed to dispersion

in the interstellar medium which is ionized hydrogen with an electron density $N_e \approx 10^5/\text{m}^3$.

(a) Show that, if $\omega \gg \omega_p$, a plot of the time delay Δt as a function of

$$\left[\frac{1}{f^2} - \frac{1}{(f+\Delta f)^2}\right]$$

is a straight line with a slope that is a measure of the distance to the pulsar.

(b) In the case of pulsar CP 0328, arrival times measured at 151, 408, and 610 MHz give the results shown below.

f Megahertz	Δt Seconds
151	
	4.18
408	
	0.367
610	

Show that, according to these measurements, the distance to CP 0328 is 266 parsec. (The *parsec* is 3.086×10^{16} meters. It is the distance from which the radius of the Earth's orbit, 1.496×10^{11} meters, would subtend an angle of one second, i.e., 1 parsec $= 1.496 \times 10^{11} \times 180 \times 3600/\pi$).

Problem P2.11

Show that any elliptically polarized wave can be decomposed into a right-hand circularly polarized wave and a left-hand circularly polarized wave.

Problem P2.12

Taking the scalar product of $\overline{D}(\bar{r},t)$ and $\partial\overline{D}(\bar{r},t)/\partial t$, show that the maximum and minimum values of $\overline{D}(\bar{r},t)$ occur at times

$$\tan 2\omega t = \frac{2\overline{D}_r \cdot \overline{D}_i}{|\overline{D}_r|^2 - |D_i|^2}$$

This yields the major and the minor axes of the ellipse. Show that the angular speed of rotation is greatest at the position of the minor axis and smallest at the position of the major axis.

Letting $\overline{D}(\bar{r},t) = \hat{x}3 + \hat{y}(1+i)$, find the major and the minor axes of the polarization ellipse and sketch the ellipse.

Problem P2.13

Show that in a uniaxial crystal the ray directions for ordinary and extraordinary rays make an angle α such that

$$\tan\alpha = \frac{(\kappa - \kappa_z)\sin\theta\cos\theta}{\kappa\cos^2\theta + \kappa_z\sin^2\theta}$$

By differentiating with respect to θ, show that maximum α occurs at

$$\tan\alpha_{max} = \frac{n_e^2 - n_o^2}{2n_o n_e}$$

where $n_e^2 = c^2\mu/\kappa_z$ and $n_o^2 = c^2\mu/\kappa$.

Problem P2.14

Consider a conductive uniaxial medium with

$$\overline{\overline{\epsilon}} = \begin{pmatrix} \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon_z \end{pmatrix}$$

and

$$\overline{\overline{\sigma}} = \begin{pmatrix} \sigma & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & \sigma_z \end{pmatrix}$$

Find dispersion relations for this medium. Explain the operation of a polaroid with this model by assuming $\sigma_z/\sigma \ll 1$. Show that a piece of polaroid turns any wave into a linearly polarized wave.

Problem P2.15

The *Fresnel ellipsoid* is defined for an anisotropic medium by

$$\epsilon_{ij}x_i x_j = 1$$

where ϵ_{ij} is expressed in the principal coordinates. The inverse of the permittivity tensor $\overline{\overline{\epsilon}}$ is $\overline{\overline{\kappa}}$, which is called the impermittivity tensor.

If we define an ellipsoid in terms of $\overline{\overline{\kappa}}$ instead of $\overline{\overline{\epsilon}}$ in the principal coordinate system of the medium and write $\kappa_{ij}x_ix_j = 1$, we have a tensor ellipsoid. Construct the Fresnel ellipsoid and the tensor ellipsoid for a uniaxial medium and a biaxial medium. Expressed in the principal coordinate system, the principal refractive indices are usually used in these definitions by replacing ϵ_{ij} with $n_i^2\delta_{ij}$ and κ_{ij} by δ_{ij}/n_i^2, in which case the tensor ellipsoid is also called an index ellipsoid or a reciprocal ellipsoid.

Problem P2.16

The electric field energy in an anisotropic crystal with $\kappa_{ij} = \kappa_{ji}$ is $W_e = \kappa_{ij}D_iD_j/2$. Writing in the principal coordinate system, show that $D_x^2/n_x^2 + D_y^2/n_y^2 + D_z^2/n_z^2 = 2W_e$. The above is identical to the defining equation of the index ellipsoid or the optical indicatrix. The index ellipsoid is useful in finding the two indices of refraction and the two corresponding directions of the $\overline{D}$ vector associated with the two characteristic waves. Find the intersection ellipse between a plane through the origin of the index ellipsoid with the surface normal along the direction of $\overline{k}$. Let $\overline{k}$ be in the $\hat{e}_3$ direction, show that the two principal axes of the ellipse are in the directions of the $\overline{D}$ vectors of the two characteristic waves.

Problem P2.17

The Fresnel equations are obtained by replacing $\overline{k}$ with an index vector $\overline{n}$ defined by $\overline{n} = \overline{k}/\omega\sqrt{\mu_o\epsilon_o}$. Discuss the dispersion relations in terms of the $\overline{n}$ vector and relate to the index ellipsoids and Fresnel ellipsoids.

Problem P2.18

Fermat's principle in geometrical optics is a variational principle, which states that the eikonal, defined by $\int \overline{k} \cdot d\overline{l}$, is a minimum for an actual optical ray between any two points. Use this principle and the dispersion relation $\overline{k} = \overline{n}(\omega/c)$ to show that the rays in an inhomogeneous medium are bent in the direction of increasing index.

Problem P2.19

Determine the dispersion relations for a uniaxial medium with

$$\overline{\overline{\kappa}} = \begin{pmatrix} \kappa & 0 & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & \kappa_z \end{pmatrix}$$

and

$$\overline{\overline{\nu}} = \begin{pmatrix} \nu & 0 & 0 \\ 0 & \nu & 0 \\ 0 & 0 & \nu_z \end{pmatrix}$$

and discuss your results.

Problem P2.20

In a ferrite, the magnetic moment $\overline{M}$ roughly obeys the relationship $d\overline{M}/dt = g\mu_o \overline{M} \times \overline{H}$, where g is the gyromagnetic ratio. When a $\hat{z}$-directed dc magnetic field $\overline{H}_0$ (zeroth order) is present, the total fields take the form $\overline{H} = \hat{z}H_0 + \overline{H}_1$, $\overline{M} = \hat{z}M_0 + \overline{M}_1$, and $\overline{B} = \mu_o(\overline{H} + \overline{M})$. Find dispersion relations for the first-order fields. Show that Faraday rotation exists in the ferrite.

Problem P2.21

Use the *kDB* system to determine the dispersion relations for a biisotropic medium. Discuss your results.

Problem P2.22

In biaxial media, the three principal dielectric constants are different. In the principal coordinate system,

$$\overline{\overline{\kappa}} = \begin{pmatrix} \kappa_x & 0 & 0 \\ 0 & \kappa_y & 0 \\ 0 & 0 & \kappa_z \end{pmatrix}$$

$$\overline{\overline{\nu}} = \nu \overline{\overline{I}}$$

$$\overline{\overline{\chi}} = \overline{\overline{\gamma}} = 0$$

The $\overline{\overline{\kappa}}$ matrix is also called the impermittivity tensor. To relate to the permittivities, we note that $\kappa_x = 1/\epsilon_x$, $\kappa_y = 1/\epsilon_y$, and $\kappa_z = 1/\epsilon_z$. The permeability ν is the reciprocal of μ.

(a) Show that the constitutive parameters in the *kDB* system take the following values:

$$\kappa_{11} = \kappa_x \sin^2\phi + \kappa_y \cos^2\phi$$
$$\kappa_{12} = \kappa_{21} = (\kappa_x - \kappa_y)\cos\theta \sin\phi \cos\phi$$
$$\kappa_{22} = (\kappa_x \cos^2\phi + \kappa_y \sin^2\phi)\cos^2\theta + \kappa_z \sin^2\theta$$
$$\kappa_{13} = \kappa_{31} = (\kappa_x - \kappa_y)\sin\theta \sin\phi \cos\phi$$
$$\kappa_{23} = \kappa_{32} = (\kappa_x \cos^2\phi + \kappa_y \sin^2\phi - \kappa_z)\sin\theta \cos\theta$$
$$\kappa_{33} = (\kappa_x \cos^2\theta + \kappa_y \sin^2\theta)\sin^2\theta + \kappa_z \cos^2\theta$$

(b) Show that the phase velocities of the characteristic waves are

$$u^2 = \frac{\nu}{2}\left[(\kappa_{11} + \kappa_{22}) \pm \sqrt{(\kappa_{11} - \kappa_{22})^2 + 4\kappa_{12}^2}\right]$$

Show that on the *kDB* plane, expressed in terms of the *kDB* base vectors $\hat{e}_1$ and $\hat{e}_2$,

$$\frac{D_2}{D_1} = \frac{\nu\kappa_{12}}{u^2 - \nu\kappa_{22}} = \frac{2\kappa_{12}}{\kappa_{11} - \kappa_{22} \pm \sqrt{(\kappa_{11} - \kappa_{22})^2 + 4\kappa_{12}^2}}$$

What are the polarizations of the two characteristic waves ? Let

$$\tan 2\psi = \frac{2\kappa_{12}}{\kappa_{11} - \kappa_{22}}$$

Show that

$$\frac{D_2}{D_1} = \tan\psi \qquad \text{or} \qquad -\cot\psi$$

Draw the two vectors on the *DB* plane. The velocities of both waves are functions of θ and ϕ. Show that none of the $\overline{E}$ vectors for the two waves lies on the *DB* plane and the $\overline{E}$ vector has a component in the $\overline{k}$ direction. Thus the energy propagation directions are different from the $\overline{k}$ direction and the two characteristic waves are both extraordinary waves.

Problem P2.23

The direction of propagation of a wave becomes ambiguous in a complex medium. From Poynting's theorem, we have learned that the

energy flow of an electromagnetic field is governed by Poynting's vector, $\overline{S} = \overline{E} \times \overline{H}$. The Poynting's vector divided by the total electromagnetic energy is refered to as the *energy velocity*. The direction of the energy velocity is thus perpendicular to both $\overline{E}$ and $\overline{H}$. We have also learned that the direction of the phase velocity is along $\overline{k}$, which is perpendicular to both $\overline{D}$ and $\overline{B}$. In a bianisotropic medium, the directions of the energy velocity and the phase velocity $\overline{k}$ do not, in general, coincide.

The Poynting power-flow direction is characterized by the ray vector $\overline{s}$. We defined the magnitude of $\overline{s}$ by

$$\overline{s} \cdot \overline{k} = 1$$

where $\overline{s}$ is perpendicular to both $\overline{E}$ and $\overline{H}$:

$$\overline{s} \cdot \overline{E} = 0$$
$$\overline{s} \cdot \overline{H} = 0$$

The ray vector $\overline{s}$ has the dimension of length.

(a) Use the vector identity $\overline{s} \times (\overline{k} \times \overline{E}) = \overline{k}(\overline{s} \cdot \overline{E}) - (\overline{k} \cdot \overline{s})\overline{E}$ to show that

$$\overline{s} \times \overline{B} = -\frac{\overline{E}}{\omega}$$
$$\overline{s} \times \overline{D} = \frac{\overline{H}}{\omega}$$

(b) Define ray surfaces similar to the wave surfaces. Show that

$$s_x^2 + s_y^2 + s_z^2 = \frac{1}{\omega^2 \mu \epsilon}$$

for the ordinary wave and

$$s_x^2 + s_y^2 + \frac{\epsilon}{\epsilon_z} s_z^2 = \frac{1}{\omega^2 \mu \epsilon_z}$$

for the extraordinary wave. Plot the ray surfaces for the negative and the positive uniaxial media.

(c) Since $\overline{s}$ is along the direction of energy velocity, $\overline{s} \cdot \delta\overline{k} = 0$; namely the normal to the wave surface gives the direction of the corresponding ray vector. Prove that the normal to the ray surface gives the direction of the corresponding $\overline{k}$ vector.

(d) The phase of a wave along a ray can be written as

$$\psi = \int \bar{k} \cdot d\bar{l} = \int \bar{k} \cdot \frac{\bar{s}}{s} dl = \frac{l}{s}$$

where l denotes the length of the segment along the ray path. In geometrical optics, the dimensionless quantity $\psi/(\omega/c)$ is the eikonal of the wave. Show that when l is equal to s multiplied by a constant, the eikonal is equal to c/ω times the constant. The ray surface gives the magnitudes of s in all directions. It follows that the ray surface describes a constant-phase surface.

Problem P2.24

Because of the presence of the Earth's magnetic field, the ionosphere becomes a gyrotropic medium. Radio wave propagation through the ionosphere is affected by Faraday rotation. The temporal variations of electron density profile in the ionosphere impose a problem on the antenna design for satellite communications if linearly polarized waves are to be used.

A linearly polarized wave at a microwave frequency f is transmitted down to Earth at an angle θ with respect to nadir, and has a small angular separation ϕ with the direction of the Earth's magnetic field $\overline{H}_e$.

(a) Assuming that the electron density N, and the Earth's magnetic field H_e are functions of height h, show that the total amount of the Faraday rotation is approximately

$$\Omega = \frac{\eta e^3 \mu_o}{8\pi^2 m^2 f^2} \int M(h) N(h)\, dh$$

where

$$\eta = \sqrt{\frac{\mu_o}{\epsilon_o}} = 377\,\Omega$$

$$M = H_e \sec\theta \cos\phi$$

and e and m are the electron charge and mass respectively.

In the above derivation, assume that the operating frequency is much higher than the plasma and cyclotron frequencies and neglect

(i) loss due to collisions between particles,

(ii) intermediate reflections due to the inhomogeneous nature of the ionosphere, and

(iii) splitting of the ordinary and the extraordinary rays.

(b) Assume that the ionosphere has a uniform electron density of $10^{11}/\mathrm{m}^3$, and the Earth's magnetic field makes a $60°$ angle with respect to nadir and has a uniform intensity of $H_e = 50$ amp/m (corresponding to a B field of 0.628 Gauss or 0.628×10^{-4} Tesla). Find the amount of the Faraday rotation for a wave of frequency 1.4 GHz transmitted from 1000 km high down to the Earth along the direction of $\overline{H}_e$.

Problem P2.25

A signal wave at frequency ω_1 can be amplified by a nonlinear medium with an intense pump wave at ω_3. This process of parametric amplification by the second-order nonlinearity generates an idler wave at $\omega_2 = \omega_3 - \omega_1$.

(a) Assuming $E_{3j}(z) = E_{3j}(0)$, show that (2.7.7) can be written as

$$\frac{dA_1}{dz} = -i\frac{\alpha}{2}A_2^* e^{-i\Delta k z}$$
$$\frac{dA_2^*}{dz} = i\frac{\alpha}{2}A_1 e^{i\Delta k z}$$

Determine α, A_1, A_2, and Δk.

(b) Let $A_2(0) = 0$ and assume that the phase-matching condition $\Delta k = 0$ is satisfied. Show that

$$A_1(z) = A_1(0)\cosh\frac{\alpha}{2}z$$
$$A_2^*(z) = iA_1(0)\sinh\frac{\alpha}{2}z$$

(c) Derive the Manley-Rowe relation

$$-\frac{\Delta P_3}{\omega_3} = \frac{\Delta P_2}{\omega_2} = \frac{\Delta P_1}{\omega_1}$$

where ΔP_l with $l = 1, 2, 3$ are the change of power between the input and output. For each photon added to the signal wave, there is a photon added to the idler wave, and a photon removed from the pump wave. Is energy conserved?

III

REFLECTION, TRANSMISSION, GUIDANCE, AND RESONANCE

3.1 Phase Matching
- a. Snell's Law
- b. Double Refraction

3.2 Reflection and Transmission at a Plane Boundary
- a. Reflection and Transmission of TE Waves
- b. Reflection and Transmission of TM Waves
- c. Wave Impedance and Transmission Line Analogy

3.3 Reflection and Transmission by a Layered Medium
- a. Reflection Coefficients
- b. Propagation Matrices and Transmission Coefficients

3.4 Guidance by Conducting Parallel Plates
- a. Transmission Line Theory
- b. Excitation of Modes in Parallel-Plate Waveguides
- c. Attenuation of Guided Waves Due to Wall Loss

3.5 Guided Waves in Layered Media
- a. Guided Waves in Isotropic-Medium-Coated Conductor
- b. Guided Waves in a Slab Dielectric Waveguide

3.6 Cylindrical Waveguides
- a. Rectangular Metallic Waveguides
- b. Circular Metallic Waveguides
- c. Circular Dielectric Waveguides

3.7 Cavity Resonators
- a. Rectangular Cavity Resonator
- b. Circular Cavity Resonator
- c. Spherical Cavity Resonator
- d. Cavity Perturbation

Problems

3.1 Phase Matching

Consider a plane boundary surface separating two homogeneous media. The boundary is at $z = 0$ [Fig. 3.1.1]. A plane wave is incident upon the boundary from the medium in region 0. A reflected wave is generated in this medium and a transmitted wave is generated in region t. The spatial dependence of the incident, the reflected, and the transmitted field vectors are, respectively, $e^{i\bar{k}_i \cdot \bar{r}}$, $e^{i\bar{k}_r \cdot \bar{r}}$, and $e^{i\bar{k}_t \cdot \bar{r}}$, where

$$\bar{k}_i = \hat{x}k_{ix} + \hat{y}k_{iy} + \hat{z}k_{iz}$$
$$= \text{wave vector for the incident wave}$$

$$\bar{k}_r = \hat{x}k_{rx} + \hat{y}k_{ry} + \hat{z}k_{rz}$$
$$= \text{wave vector for the reflected wave}$$

$$\bar{k}_t = \hat{x}k_{tx} + \hat{y}k_{ty} + \hat{z}k_{tz}$$
$$= \text{wave vector for the transmitted wave}$$

The boundary conditions at $z = 0$ require that tangential $\bar{E}$ and $\bar{H}$ fields be continuous for all x and y. Notice that the wave vectors $\bar{k}_i$, $\bar{k}_r$, and $\bar{k}_t$ indicate the directions of the propagation of the constant phase fronts where the uniform plane wave fields take constant values in amplitude and phase. Thus at any point in region 0, there are reflected wave fields as well as incident wave fields. As an example, let the magnitudes of the tangential components of the electric field vectors be E_i, E_r, and E_t. The boundary conditions at $z = 0$ then give

$$E_i e^{ik_{ix}x + ik_{iy}y} + E_r e^{ik_{rx}x + ik_{ry}y} = E_t e^{ik_{tx}x + ik_{ty}y} \tag{1}$$

This equation must hold for all x and y. The only way that this equation can be satisfied for all x and y is to have $E_i + E_r = E_t$,

$$k_{ix} = k_{rx} = k_{tx} = k_x$$

and

$$k_{iy} = k_{ry} = k_{ty} = k_y$$

Thus, as a consequence of the boundary conditions, the tangential components of all wave vectors are continuous. This result is known as the *phase-matching conditions.*

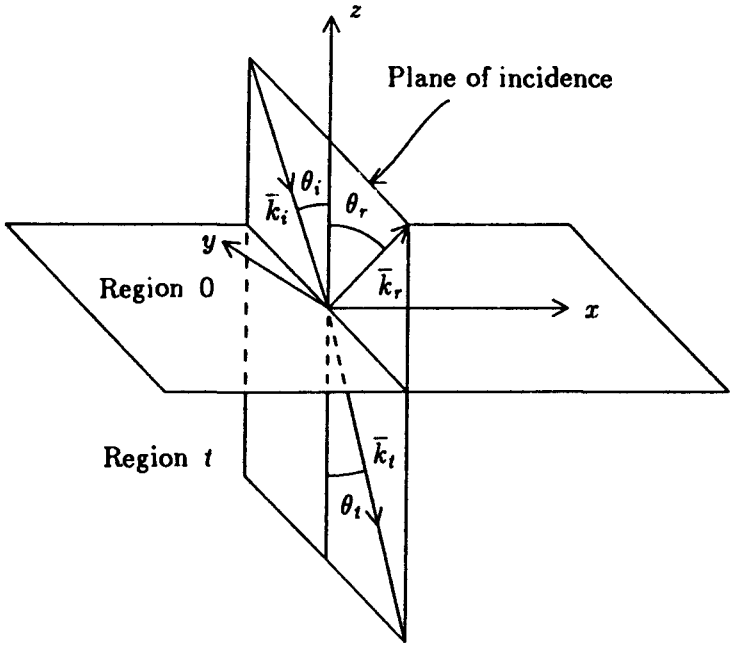


Figure 3.1.1 Wave vectors $\bar{k}_i$, $\bar{k}_r$ and $\bar{k}_t$.

a. Snell's Law

According to the phase-matching conditions, the incident, reflected and transmitted wave vectors must all lie in the same plane called the *plane of incidence* [Fig. 3.1.1]. The plane of incidence is determined by the incident $\bar{k}_i$ vector and the normal to the boundary surface. Thus at normal incidence, the plane is not defined. Consider the case when both media are isotropic. The magnitudes of the wave vectors are characterized by the dispersion relations

$$k_i^2 = k_r^2 = \omega^2 \mu_0 \epsilon_0 \tag{2}$$

and

$$k_t^2 = \omega^2 \mu_t \epsilon_t \tag{3}$$

Let θ_i denote the angle of incidence, θ_r the angle of reflection, and θ_t the angle of transmission, with θ_i, θ_r, and θ_t all less than $\pi/2$ [Fig. 3.1.1]. Since the tangential components of the three wave vectors are continuous, we find

$$k_i \sin\theta_i = k_r \sin\theta_r \tag{4}$$

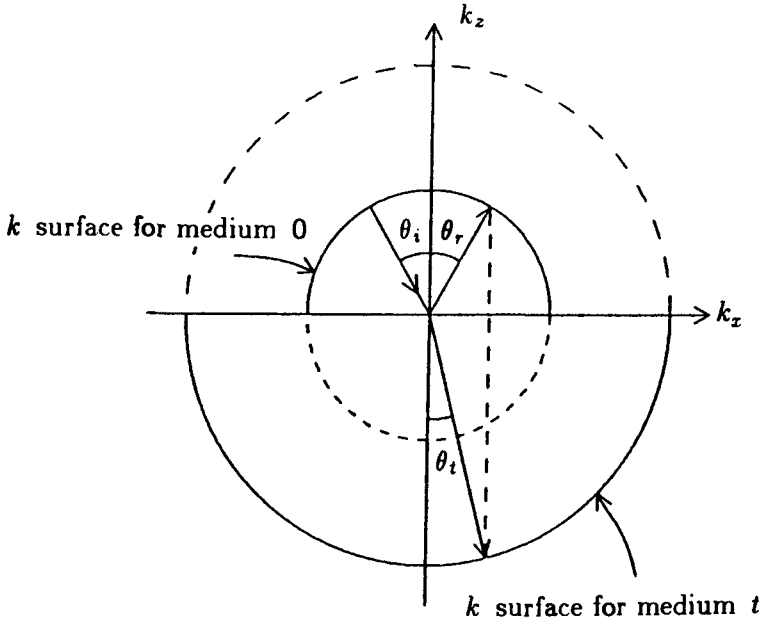


Figure 3.1.2 Phase matching using k surfaces.

$$k_i \sin\theta_i = k_t \sin\theta_t \tag{5}$$

In view of (2), (4) states that the angle of reflection θ_r is equal to the angle of incidence θ_i. By virtue of (3), (5) gives

$$\frac{\sin\theta_i}{\sin\theta_t} = \frac{n_t}{n_0} \tag{6}$$

where $n_t = c\sqrt{\mu_t\epsilon_t}$ and $n_0 = c\sqrt{\mu_0\epsilon_0}$ are the refractive indices. Equation (6) is known as Snell's law.

We may illustrate the phase-matching conditions by a simple figure [Fig. 3.1.2]. Assuming that the plane of incidence lies in the $x-z$ plane, we plot the k surfaces for the two isotropic media. Let $n_0 < n_t$ be such that the radius of the k surface for region 0 has a shorter radius than that of region t. The plot is constructed on the $k_x - k_z$ plane in k space. Although the k_x axis is not a physical axis, it provides the feeling of a physical boundary. Therefore we use dashed lines for the k surface of region 0 below the k_x line and similarly for the k

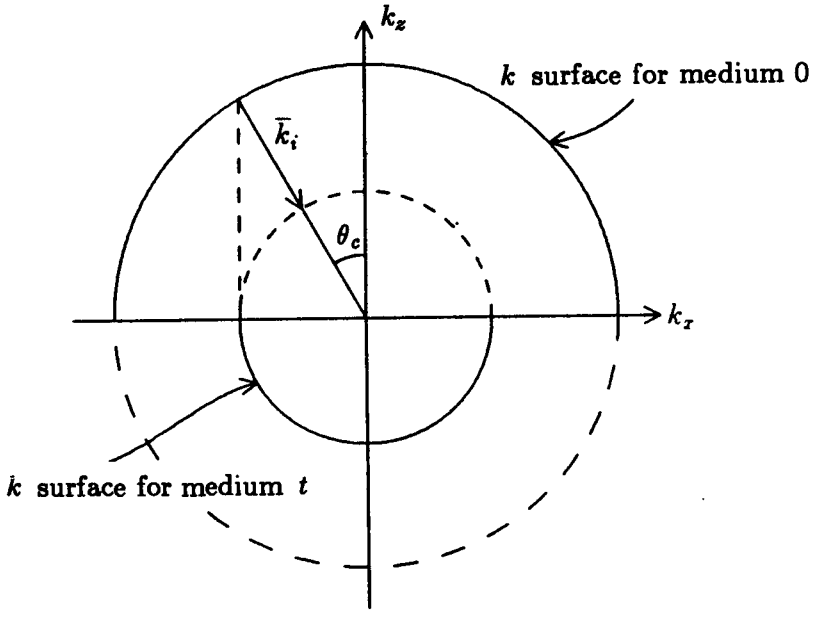


Figure 3.1.3 k **surface for medium 0 is larger than that for medium** t.

surface of region t above the k_x line. We plot the incident $\bar{k}$ vector from the k surface toward the origin. Requiring that the tangential k_x components be equal, we construct the $\bar{k}$ vectors for the reflected and transmitted waves as shown in Figure 3.1.2.

However, it is not always possible to construct the reflected $\bar{k}_r$ vector or the transmitted $\bar{k}_t$ vector. Suppose that the medium in region 0 is denser than the medium in region t such that $n_0 > n_t$. Then the radius of the k surface in region t is shorter than that in region 0 [Fig. 3.1.3]. By the phase-matching conditions, we see that as k_{ix} of the incident wave becomes larger than k_t, there is no intersection with the small circle because this amounts to requiring that one component of a vector be greater than its magnitude — an impossibility unless the vector is complex. The k surface in region t is described by

$$k_{tz}^2 + k_{ix}^2 = k_t^2 \tag{7}$$

Since $k_{ix} > k_t$, k_{tz} must be purely imaginary,

$$k_{tz} = \pm\sqrt{k_t^2 - k_{ix}^2} = \pm i\alpha_{tz} \tag{8}$$

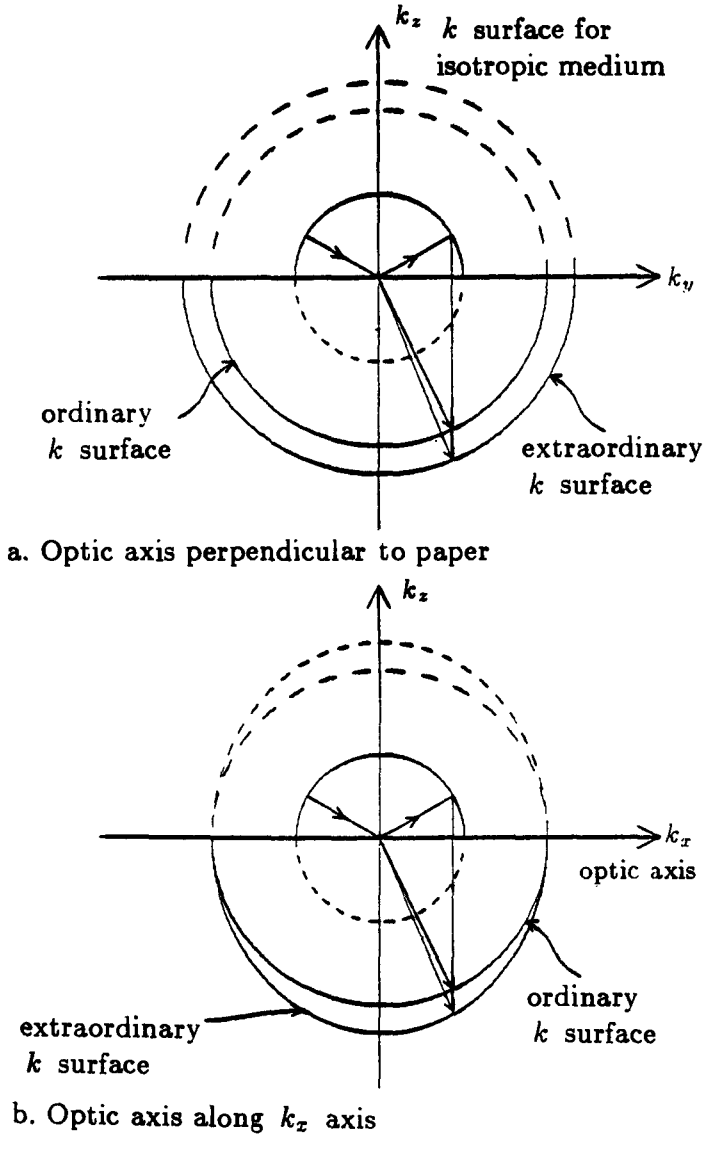


Figure 3.1.4 Double refraction by uniaxial medium.

Remember that the wave in region t is characterized by $\exp(ik_{tz}z + ik_x x)$. For $k_x > k_t$, it becomes $\exp(\alpha_{tz}z + ik_x x)$. We have chosen the minus sign in (8) such that the transmitted wave decays exponentially

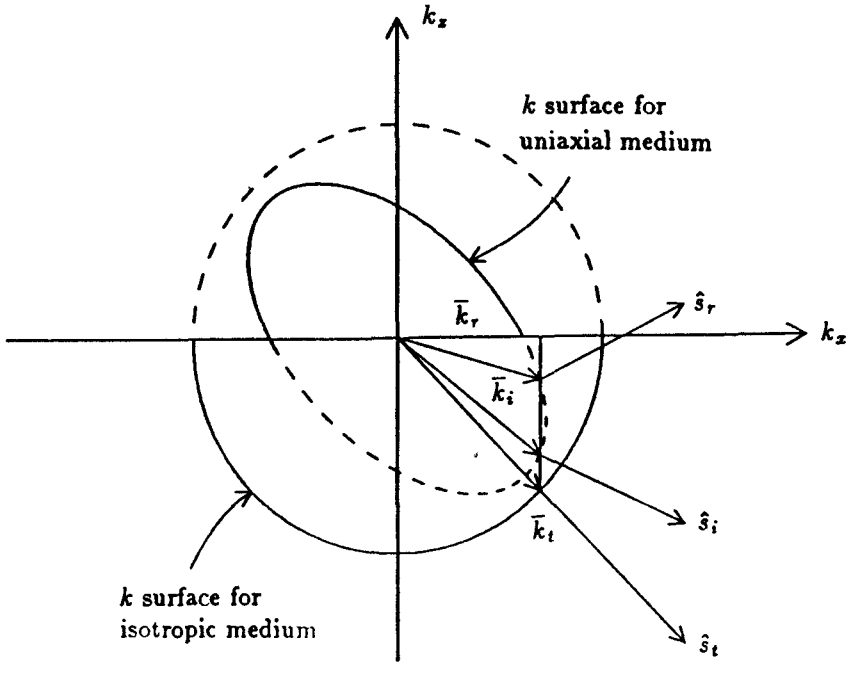


Figure 3.1.5 Wave vector directions and power flow directions.

in the negative $\hat{z}$ direction. The transmitted wave thus propagates along the $\hat{x}$ direction with the phase velocity ω/k_x. This can be regarded as a plane wave with constant phase fronts perpendicular to the boundary surface. Its amplitude is maximum at the surface and decays exponentially away from the surface. The wave is known as a surface wave. The surface wave is evanescent in the $-\hat{z}$ direction. Since evanescence in the transmitted wave begins when $k_t = k_x = k \sin\theta_c$, the angle θ_c is the critical angle of incidence. When the wave amplitudes are solved and the energy flow of the waves is determined, we shall find that there is no time-average power penetrating the medium t when the incident angle is greater than θ_c. For this reason, this phenomenon is considered to be total reflection.

b. Double Refraction

Construction of the k surface enables us to visualize the $\bar{k}$ vectors in the reflected and transmitted regions [Figs. 3.1.2-3]. This tech-

nique can also be applied to media that are not isotropic. Consider the phenomenon of double refraction by a positive uniaxial medium [Fig. 3.1.4]. First, let the optic axis of the medium be perpendicular to the plane of incidence. The two transmitted wave vectors are shown in Figure 3.1.4a. The power flow directions for the ordinary and extraordinary waves are the same as the directions of the wave vectors. Next, let the optic axis be parallel to the plane of incidence. The two transmitted wave vectors are shown in Figure 3.1.4b. By the nature of the wave surface, the power flow direction of the extraordinary wave is no longer the same as the direction of $\overline{k}$. Note that by proper source excitation we can generate either the ordinary or the extraordinary wave. For instance, if the wave is linearly polarized perpendicular to the plane of incidence only ordinary waves are excited.

Consider another wave excited in a uniaxial medium and incident upon the interface of an isotropic medium. Let the optic axis be in the plane of incidence and make an angle with the boundary. The wave surfaces are shown in Figure 3.1.5. In Figures 3.1.2–4, we have drawn the incident $\overline{k}_i$ vector with arrows pointing toward the origin. Although we use dotted lines for half of the spheres to convey the feeling of a physical boundary, we must realize that the plot is for k space and not physical space. In Figure 3.1.5 we draw the reflected wave vector; instead of pointing in the positive k_x and k_z directions, it is now pointing in the positive k_x and negative k_z directions. The power flow direction for the reflected wave, however, is pointing in the positive k_x and k_z directions. Thus the reflected wave, while carrying energy away from the interface, has its phase front propagating toward the interface. This is called a *backward wave* with respect to the normal at the interface.

3.2 Reflection and Transmission at a Plane Boundary

An incident wave with any polarization can be decomposed into TE (transverse electric) and TM (transverse magnetic) wave components. The TE wave is linearly polarized with the electric field vector perpendicular to the plane of incidence and is called perpendicularly polarized, horizontally polarized, or simply an $\overline{E}$ wave or s wave. The TM wave is linearly polarized with the electric vector parallel to the plane of incidence and is called parallelly polarized, vertically polarized, or simply an $\overline{H}$ wave or p wave. The two wave components can

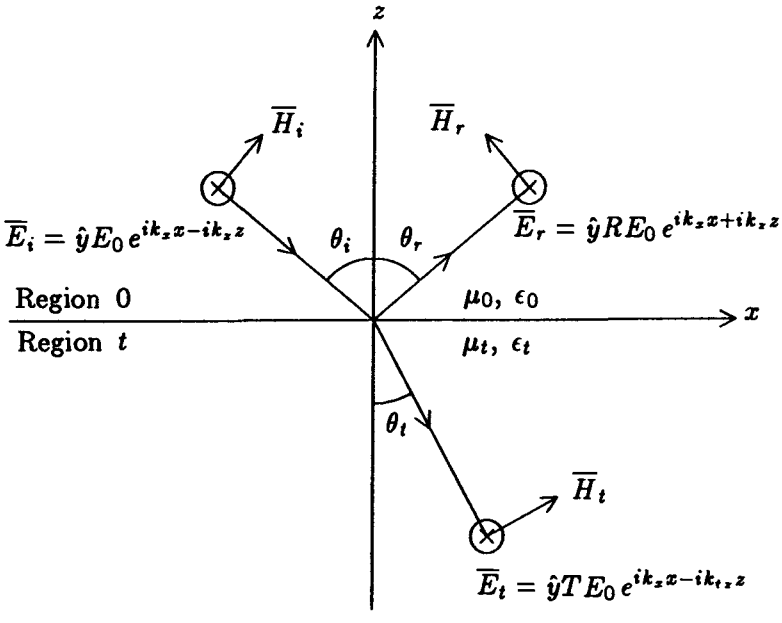


Figure 3.2.1 Reflection and transmission of TE waves at a plane boundary separating regions 0 and t.

be studied separately.

Consider a plane wave incident from an isotropic medium with permittivity ϵ_0 and permeability μ_0 upon another isotropic medium with permittivity ϵ_t and permeability μ_t [Fig. 3.2.1]. We assume the plane of incidence to be parallel to the $x-z$ plane. By the phase-matching conditions the $\overline{k}$ vectors of all plane waves will be in the $x-z$ plane. Thus all field vectors will be dependent on x and z only and independent of y. Since $\partial/\partial y = 0$, the Maxwell equations become, in region 0,

$$H_x = -\frac{1}{i\omega\mu_0}\frac{\partial}{\partial z}E_y \tag{1}$$

$$H_z = \frac{1}{i\omega\mu_0}\frac{\partial}{\partial x}E_y \tag{2}$$

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} + \omega^2\mu_0\epsilon_0\right)E_y = 0 \tag{3}$$

$$E_x = \frac{1}{i\omega\epsilon_0}\frac{\partial}{\partial z}H_y \tag{4}$$

$$E_z = -\frac{1}{i\omega\epsilon_0}\frac{\partial}{\partial x}H_y \tag{5}$$

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} + \omega^2\mu_0\epsilon_0\right)H_y = 0 \tag{6}$$

Equations (1) and (2) are the x and z components of Faraday's law and (4) and (5) are the x and z components of source-free Ampere's law. Equation (3) is obtained from the y component of the source-free Ampere's law by making use of (1)–(2), and (6) is obtained from the y component of Faraday's law by making use of (4) and (5). A similar set of equations hold in region t with μ_0 and ϵ_0 replaced by μ_t and ϵ_t. Equations (3) and (6) are called homogeneous Helmholtz wave equations. We see that TE waves are completely determined by (1)–(3), which are *decoupled* from (4)–(6) governing TM waves. The two sets of equations are dual of each other. We obtain (1)–(3) from (4)–(6) or vice versa by the replacements $\overline{E} \to \overline{H}$, $\overline{H} \to -\overline{E}$, and $\mu_0 \rightleftarrows \epsilon_0$. This is known as the principle of duality.

a. Reflection and Transmission of TE Waves

For a TE plane wave incident on the isotropic medium, the total electric and magnetic field vectors in region 0 are the sum of the reflected and incident components [Fig. 3.2.1]. We write

$$E_y = (RE_0\, e^{ik_z z} + E_0\, e^{-ik_z z})e^{ik_x x} \tag{7}$$

$$H_x = -\frac{k_z}{\omega\mu_0}(RE_0\, e^{ik_z z} - E_0\, e^{-ik_z z})e^{ik_x x} \tag{8}$$

$$H_z = \frac{k_x}{\omega\mu_0}(RE_0\, e^{ik_z z} + E_0\, e^{-ik_z z})e^{ik_x x} \tag{9}$$

where E_0 is the incident wave amplitude and R denotes the reflection coefficient for electric fields. Equations (8) and (9) are obtained from (7) by using (1) and (2). In the exponents of (7)–(9), the x components of the $\overline{k}$ vector for the incident and reflected waves are equal by virtue of phase matching. The z components are opposite in sign, indicating that the reflected and incident waves are propagating along opposite $\hat{z}$ directions. The magnitudes are equal because they obey the same dispersion relation

$$k_x^2 + k_z^2 = \omega^2\mu_0\epsilon_0 \tag{10}$$

This is easily observed by substituting (7) in the Helmholtz wave equation (3).

In region t, there is only a transmitted wave [Fig. 3.2.1]. The TE wave solutions take the form

$$E_{ty} = TE_0\, e^{-ik_{tz}z + ik_x x} \tag{11}$$

$$H_{tx} = \frac{k_{tz}}{\omega\mu_t} TE_0\, e^{-ik_{tz}z + ik_x x} \tag{12}$$

$$H_{tz} = \frac{k_x}{\omega\mu_t} TE_0\, e^{-ik_{tz}z + ik_x x} \tag{13}$$

where T denotes the transmission coefficient for the electric field. Again this set of equations satisfies (1)–(3) and the dispersion relation

$$k_x^2 + k_{tz}^2 = \omega^2 \mu_t \epsilon_t \tag{14}$$

The x component of $\overline{k}_t$ is equal to those of the incident and reflected waves on account of phase matching.

Let the boundary surface be at $z = 0$ where the tangential components of $\overline{E}$ and $\overline{H}$ are continuous. We obtain

$$R + 1 = T \tag{15}$$

$$\frac{k_z}{\mu_0}(R - 1) = -\frac{k_{tz}}{\mu_t} T \tag{16}$$

Note that we did not use the boundary conditions that normal $\overline{D}$ and normal $\overline{B}$ components are continuous at $z = 0$, because these two conditions are not independent of the two tangential $\overline{E}$ and $\overline{H}$ conditions, just as Gauss' two laws are not independent of Faraday's and Ampere's laws. In this case we can see that the condition of continuous normal $\overline{B}$ yields the same equation as (15) and there is no normal $\overline{D}$ component.

The reflection and transmission coefficients R and T are easily determined from (15) and (16),

$$R = R_{0t} = \frac{1 - p_{0t}}{1 + p_{0t}} \tag{17}$$

and

$$T = T_{0t} = \frac{2}{1 + p_{0t}} \tag{18}$$

where, for TE waves

$$p_{0t} = \frac{\mu_0 k_{tz}}{\mu_t k_z} \tag{19}$$

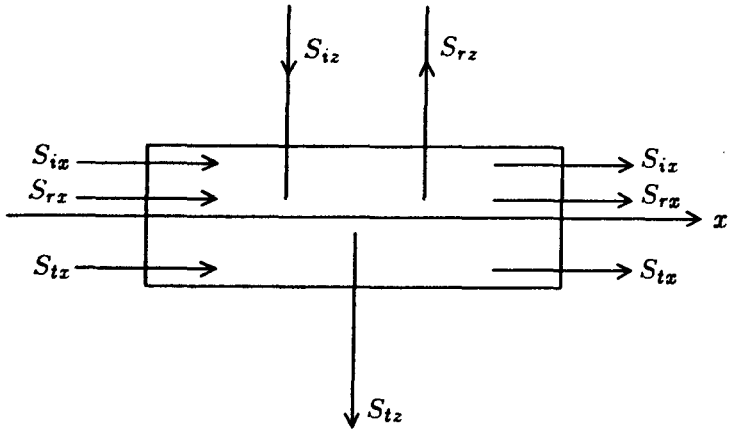


Figure 3.2.2 Power conservation at a plane boundary.

With p_{0t} for TE waves defined in (19), R_{0t} in (17) is called the Fresnel reflection coefficient for a TE wave incident from region 0 and reflected at the boundary separating regions 0 and t. In (18) T_{0t} is the transmission coefficient from region 0 to region t.

It is easily shown that when the boundary surface is not at $z = 0$ but at $z = -d_0$, then the phases of the reflected and transmitted waves are dependent on the location of the boundary surface. The reflection coefficient R gains a phase shift of $2k_z d_0$ and the transmission coefficient gains a phase shift of $(k_z - k_{tz})d_0$.

With the reflection and transmission coefficients known, the time-average Poynting vectors for the incident, the reflected, and the transmitted waves are calculated to be

$$< \overline{S}_i > = \frac{1}{2}\left(-\hat{z}\frac{k_z}{\omega\mu_0} + \hat{x}\frac{k_x}{\omega\mu_0}\right)|E_0|^2 \tag{20}$$

$$< \overline{S}_r > = \frac{1}{2}\left(\hat{z}\frac{k_z}{\omega\mu_0} + \hat{x}\frac{k_x}{\omega\mu_0}\right)|R|^2\,|E_0|^2 \tag{21}$$

$$< \overline{S}_t > = \frac{1}{2}\left(-\hat{z}\frac{k_{tz}}{\omega\mu_t} + \hat{x}\frac{k_x}{\omega\mu_t}\right)|T|^2\,|E_0|^2 \tag{22}$$

where we have assumed that both k_x and k_{tz} are real. Power conservation is observed by considering a control volume across the boundary [Fig. 3.2.2]. We see that the x components of all the Poynting vectors

entering and exiting the control volume are equal. In order to have power conservation we must prove that the sum of the normal components of the reflected power and the transmitted power is equal to the normal component of the incident power. We define the power reflection coefficient or the *reflectivity* to be

$$r = \frac{\hat{z} \cdot < \overline{S}_r >}{-\hat{z} \cdot < \overline{S}_i >} = |R|^2 \tag{23}$$

By the same token, we define the power transmission coefficient or the *transmissivity* to be

$$t = \frac{-\hat{z} \cdot < \overline{S}_t >}{-\hat{z} \cdot < \overline{S}_i >} = p_{0t} |T|^2 \tag{24}$$

By virtue of (17)–(19), we see that $r+t = 1$. This demonstrates power conservation for reflection and transmission at a boundary surface.

When the incident angle is larger than the critical angle, k_{tz} becomes imaginary. In view of (11)–(13) and noticing that z is negative inside region t, we write $k_{tz} = i\alpha_{tz}$ where α_{tz} is positive real. The transmitted complex Poynting vector becomes

$$\overline{S}_t = \left(-\hat{z} \frac{i\alpha_{tz}}{\omega \mu_t} + \hat{x} \frac{k_x}{\omega \mu_t} \right) e^{2\alpha_{tz}(z+d_0)} |T|^2 |E_0|^2 \tag{25}$$

The real part of the z component of $\overline{S}_t$ is zero and we see that there is no time-average power transmitted in the $\hat{z}$ direction into region t. The transmissivity t is thus zero and the reflectivity r must be unity. This is seen from the Fresnel reflection coefficient which now reads

$$R_{0t} = \frac{1 - p_{0t}}{1 + p_{0t}} = e^{i2\phi} \tag{26}$$

where

$$p_{0t} = i \frac{\mu_0 \alpha_{tz}}{\mu_t k_z} \tag{27}$$

for TE waves and

$$\phi = -\tan^{-1} \frac{\mu_0 \alpha_{tz}}{\mu_t k_z} \tag{28}$$

The magnitude of R_{0t} is unity. Thus all the time-average incident power in the $\hat{z}$ direction is reflected. The phase shift ϕ of the Fresnel

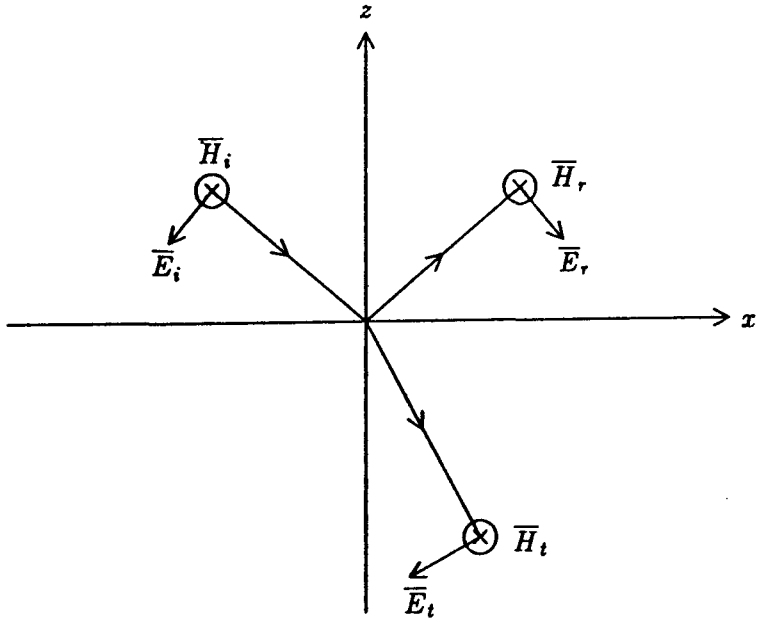


Figure 3.2.3 Reflection and transmission of TM waves.

reflection coefficient at total reflection is known as the Goos–Hänchen shift.

b. Reflection and Transmission of TM Waves

The reflection and transmission of TM waves by a plane boundary can be carried out in a manner similar to the treatment of TE waves. We can also invoke the principle of duality and write down the answers directly. Making the replacement $\overline{E} \to \overline{H}$, $\overline{H} \to -\overline{E}$, $\mu_0 \rightleftarrows \epsilon_0$, and the boundary conditions of continuous tangential $\overline{H}$ and $\overline{E}$ at $z = 0$, we find the dual of the TE problem to be precisely the TM problem [Fig. 3.2.3]. We obtain the reflection and transmission coefficients as in (17)–(18) with p_{0t} in (19) replaced by

$$p_{0t} = \frac{\epsilon_0 k_{tz}}{\epsilon_t k_z} \tag{29}$$

for TM waves. The reflectivity and transmissivity are the same as (23) and (24). The Goos–Hänchen phase shift at total internal reflection is,

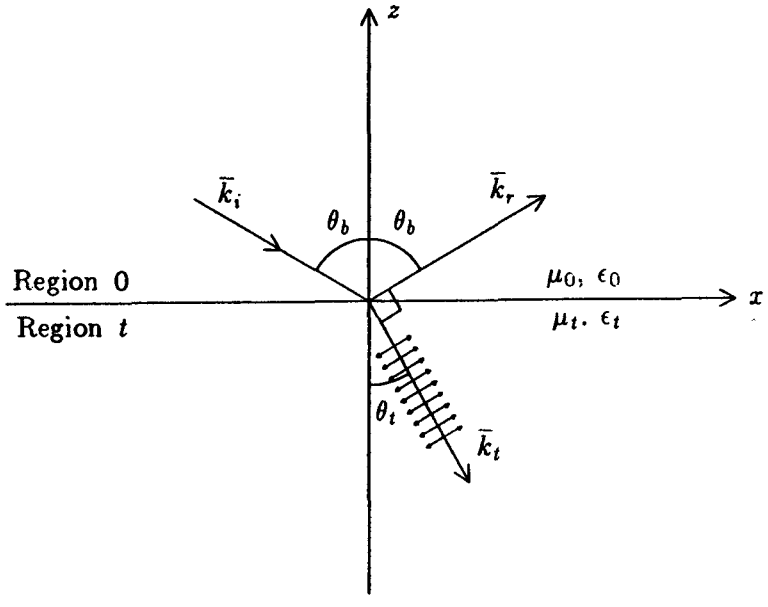


Figure 3.2.4 Incidence at the Brewster angle.

is, instead of (28),

$$\phi = -\tan^{-1}\frac{\epsilon_0 \alpha_{tz}}{\epsilon_t k_z} \tag{30}$$

for TM waves. Note that the Fresnel reflection coefficient for TM waves is now representing the ratio of the reflected and incident magnetic fields.

At the surface of a perfect conductor, we may calculate the reflection coefficients by letting $\epsilon_t \to \infty$. We find for TE waves $p_{0t} \to \infty$ and $R_{0t} \to -1$ while for TM waves $p_{0t} \to 0$ and $R_{0t} \to 1$. Thus the tangential electric field vanishes at the boundary and the tangential magnetic field doubles its strength in order to support the induced surface currents.

We now discuss the Brewster angle for nonmagnetic media where $\mu_t = \mu_0$. The Brewster angle θ_b is the incident angle at which there is no reflected power. Setting $R_{0t} = 0$ or $p_{0t} = 1$ we find, from (19),

$$k_{tz} = k_z \tag{31}$$

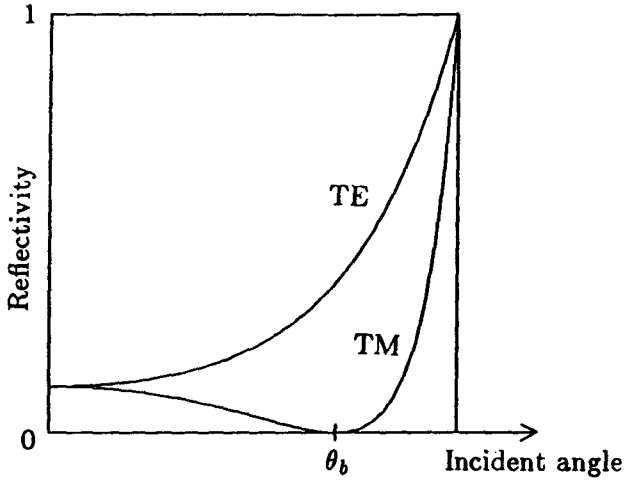


Figure 3.2.5 Reflectivity of TE and TM waves.

for TE waves, and from (29),

$$\epsilon_0 k_{tz} = \epsilon_t k_z \tag{32}$$

for TM waves. In view of the fact that $k_{tz} = k_z$ only if both dielectrics have identical permittivities, we see that no Brewster angle can exist for the TE waves. For TM waves, (32) can be written as

$$k \cos\theta_t = k_t \cos\theta_b \tag{33}$$

where we use the fact that $\epsilon_0/\epsilon_t = k^2/k_t^2$ and that $k_z = k\cos\theta_b$ with θ_b denoting the Brewster angle of incidence. Also, by Snell's law,

$$k \sin\theta_b = k_t \sin\theta_t \tag{34}$$

Cancelling k and k_t from (33) and (34) we find $\sin 2\theta_b = \sin 2\theta_t$. The solution $\theta_t = \theta_b$ again implies identical media and is of no interest. In view of the fact that both θ_b and θ_t are larger than zero and less than 90°, we obtain as the other solution,

$$\theta_b + \theta_t = \frac{\pi}{2} \tag{35}$$

Since the reflected direction is perpendicular to the transmitted direction, the reflected wave vector $\overline{k}_r$ is perpendicular to the transmitted wave vector $\overline{k}_t$ [Fig. 3.2.4].

Physically we can explain this by visualizing the dielectric media as consisting of dipoles that are excited by the transmitted wave and radiating at the same frequency. Each individual dipole has a radiation pattern that is maximum in a direction perpendicular to the dipole axis and null along the dipole axis. For a TM wave excitation, all dipoles oscillate parallel to the plane of incidence along the $\overline{E}$-field lines. At the Brewster angle of incidence, the reflected $\overline{k}_r$ vector is in the same direction as the dipole oscillation in the transmitted medium. Thus, no TM wave is reflected.

Substituting (35) in (34), we obtain the Brewster angle

$$\theta_b = \tan^{-1}\frac{k_t}{k} = \tan^{-1}\sqrt{\frac{\epsilon_t}{\epsilon_0}} \tag{36}$$

In Figure 3.2.5, we plot the reflectivities as functions of incident angles. In general, on a solid dielectric surface, the TE waves reflect more than the TM waves. When an unpolarized wave is incident at θ_b upon an isotropic medium, the reflected wave becomes linearly polarized perpendicular to the plane of incidence. Thus the Brewster angle is also referred to as the polarization angle.

c. Wave Impedance and Transmission Line Analogy

We can define a wave impedance $Z(z)$ in the transmitted region by taking the ratio of (11) and (12) which gives

$$Z_t = \frac{E_{ty}}{H_{tx}} = \frac{\omega\mu_t}{k_{tz}} \tag{37}$$

At normal incidence, $k_{tz} = \omega(\mu_t\epsilon_t)^{1/2}$ and the wave impedance is simply the characteristic impedance of medium t, $Z_t = (\mu_t/\epsilon_t)^{1/2}$. At an incident angle larger than θ_c, Z_t is purely imaginary.

The wave impedance in region 0 is defined by the ratio of (7) to (8), which gives

$$Z(z) = \frac{E_y}{H_x} = \frac{\omega\mu_o}{k_z}\frac{1 + Re^{i2k_z z}}{1 - Re^{i2k_z z}}$$

For a perfect conductor occupying region t, $R = -1$. We observe that as $2k_z z$ varies from 0 to 2π, the wave impedance repeats its

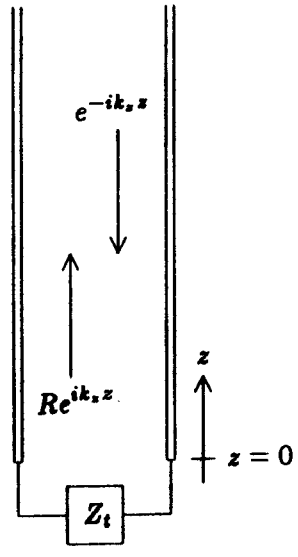


Figure 3.2.6 Transmission line model.

value at $z = 0$. This is also related to the periodic variation of the field amplitudes in the z direction. As seen from (7),

$$E_y = E_0 e^{ik_x x - ik_z z}(1 + Re^{i2k_z z})$$

the expression in the parentheses repeats its values over every interval of 2π for $2k_z z$. This phenomenon is best described in terms of the transmission line theory to be described in 3.4. The transmission line model for (7) and (8) is shown in Figure 3.2.6, where Z_t is the wave impedance for region t.

3.3 Reflection and Transmission by a Layered Medium

Consider a plane wave incident on a stratified isotropic medium with boundaries at $x = -d_0, -d_1, \ldots, -d_n$ [Fig. 3.3.1]. The $(n+1)$ th region is semi-infinite and is labeled region t, $t = n+1$. The permittivity and permeability in each region are denoted by ϵ_l and μ_l. The plane wave is incident from region 0 and has the plane of incidence parallel to the $x - z$ plane. All field vectors are dependent on x and z only and independent of y. Since $\partial/\partial y = 0$, the

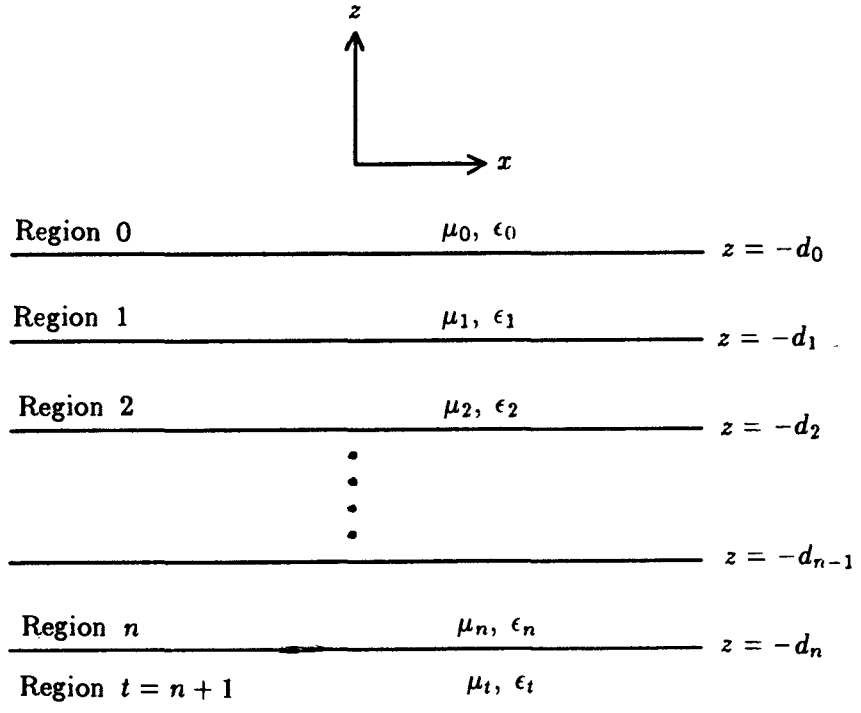


Figure 3.3.1 Layered medium.

Maxwell equations in any region l can be separated into TE and TM components governed by E_{ly} and H_{ly}. We obtain

$$H_{lx} = -\frac{1}{i\omega\mu_l}\frac{\partial}{\partial z}E_{ly} \tag{1}$$

$$H_{lz} = \frac{1}{i\omega\mu_l}\frac{\partial}{\partial x}E_{ly} \tag{2}$$

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} + \omega^2\mu_l\epsilon_l\right)E_{ly} = 0 \tag{3}$$

$$E_{lx} = \frac{1}{i\omega\epsilon_l}\frac{\partial}{\partial z}H_{ly} \tag{4}$$

$$E_{lz} = -\frac{1}{i\omega\epsilon_l}\frac{\partial}{\partial x}H_{ly} \tag{5}$$

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} + \omega^2\mu_l\epsilon_l\right)H_{ly} = 0 \tag{6}$$

The TE waves are completely determined by (1)–(3) and the TM waves by (4)–(6). The two sets of equations are duals of each other under the replacements $\overline{E}_l \to \overline{H}_l$, $\overline{H}_l \to -\overline{E}_l$, and $\mu_l \rightleftarrows \epsilon_l$.

For a TE plane wave, $E_y = E_0 e^{-ik_{lz}z + ik_x x}$, incident on the stratified medium, the total field in region l can be written as

$$E_{ly} = \left(A_l e^{ik_{lz}z} + B_l e^{-ik_{lz}z}\right) e^{ik_x x} \tag{7}$$

$$H_{lx} = -\frac{k_{lz}}{\omega\mu_l}\left(A_l e^{ik_{lz}z} - B_l e^{-ik_{lz}z}\right) e^{ik_x x} \tag{8}$$

$$H_{lz} = \frac{k_x}{\omega\mu_l}\left(A_l e^{ik_{lz}z} + B_l e^{-ik_{lz}z}\right) e^{ik_x x} \tag{9}$$

Obviously (7) satisfies the Helmholtz wave equation in (3). Substitution of (7) in (3) yields the dispersion relation

$$k_{lz}^2 + k_x^2 = \omega^2 \mu_l \epsilon_l \tag{10}$$

We do not write a subscript l for the x component of $\overline{k}$ as a consequence of the phase-matching conditions. Truly, there are multiple reflections and transmissions in each layer l. The amplitude A_l thus represents all wave components that have a propagating velocity component along the positive $\hat{z}$ direction, and B_l represents those with a velocity component along the negative $\hat{z}$ direction.

We note that in region 0 where $l = 0$,

$$A_0 = RE_0 \tag{11}$$

$$B_0 = E_0 \tag{12}$$

In region t where $l = n + 1 = t$, we have

$$A_t = 0 \tag{13}$$

$$B_t = TE_0 \tag{14}$$

because region t is semi-infinite and there is no wave propagating with a velocity component in the positive $\hat{z}$ direction. We denote the transmitted amplitude by T.

The wave amplitudes A_l and B_l are related to wave amplitudes in neighboring regions by the boundary conditions. At $z = -d_l$, boundary conditions require that E_y and H_x be continuous. We obtain

$$A_l e^{-ik_{lz}d_l} + B_l e^{ik_{lz}d_l} = A_{l+1} e^{-ik_{(l+1)z}d_l} + B_{l+1} e^{ik_{(l+1)z}d_l} \tag{15}$$

$$\frac{k_{lz}}{\mu_l}\left[A_l e^{-ik_{lz}d_l} - B_l e^{ik_{lz}d_l}\right] = \frac{k_{(l+1)z}}{\mu_{l+1}}\left[A_{l+1}e^{-ik_{(l+1)z}d_l} - B_{l+1}e^{ik_{(l+1)z}d_l}\right] \tag{16}$$

There are $n+1$ boundaries which give rise to $(2n+2)$ equations. In region 0, we have an unknown reflection coefficient R. In region t, we have an unknown transmission coefficient T. There are two unknowns A_l and B_l in each of the regions $l = 1, 2, \ldots, n$. Thus we have a total of $(2n+2)$ unknowns. To solve for the $(2n+2)$ unknowns from the $(2n+2)$ linear equations, we can arrange the equations in matrix form with the unknowns forming a $(2n+2)$ column matrix and the coefficients forming a $(2n+2) \times (2n+2)$ square matrix. The solution is then obtained by inverting the square matrix. This procedure is straightforward but tedious. We shall now describe simpler ways to deal with the problem.

a. Reflection Coefficients

Since we are interested in finding the reflection coefficient for the stratified medium, let us derive a closed-form formula for R. We first solve (15) and (16) for A_l and B_l.

$$A_l e^{-ik_{lz}d_l} = \frac{1}{2}\left(1 + p_{l(l+1)}\right)\left\{A_{l+1}e^{-ik_{(l+1)z}d_l} + R_{l(l+1)}B_{l+1}e^{ik_{(l+1)z}d_l}\right\} \tag{17}$$

$$B_l e^{ik_{lz}d_l} = \frac{1}{2}\left(1 + p_{l(l+1)}\right)\left\{R_{l(l+1)}A_{l+1}e^{-ik_{(l+1)z}d_l} + B_{l+1}e^{ik_{(l+1)z}d_l}\right\} \tag{18}$$

where

$$p_{l(l+1)} = \frac{\mu_l k_{(l+1)z}}{\mu_{l+1} k_{lz}} \tag{19}$$

for TE waves and

$$R_{l(l+1)} = \frac{1 - p_{l(l+1)}}{1 + p_{l(l+1)}} \tag{20}$$

is the reflection coefficient for waves in region l, caused by the boundary separating regions l and $l+1$. We note from (19) that

$$p_{(l+1)l} = \frac{1}{p_{l(l+1)}} \tag{21}$$

which also gives

$$R_{(l+1)l} = -R_{l(l+1)} \tag{22}$$

Thus the reflection coefficient in region $l+1$, $R_{(l+1)l}$, caused by the boundary separating regions $l+1$ and l, is equal to the negative of $R_{l(l+1)}$.

Forming the ratio of (17) and (18) we obtain

$$\begin{aligned}\frac{A_l}{B_l} &= \frac{e^{i2k_{lz}d_l}}{R_{l(l+1)}} + \frac{\left[1-\left(1/R_{l(l+1)}^2\right)\right]e^{i2(k_{(l+1)z}+k_{lz})d_l}}{\left[1/R_{l(l+1)}\right]e^{i2k_{(l+1)z}d_l} + (A_{l+1}/B_{l+1})} \\ &= \frac{e^{i2k_{lz}d_l}}{R_{l(l+1)}} + \cfrac{\left[1-\left(1/R_{l(l+1)}^2\right)\right]e^{i2(k_{(l+1)z}+k_{lz})d_l}}{\left[1/R_{l(l+1)}\right]e^{i2k_{(l+1)z}d_l}} + \frac{A_{l+1}}{B_{l+1}}\end{aligned} \tag{23}$$

With the second equality we introduce a notation for writing a continued fraction. Equation(23) expresses (A_l/B_l) in terms of A_{l+1}/B_{l+1} and so on, until the transmitted region t, where $A_t/B_t = 0$, is reached.

The reflection coefficient due to the stratified medium is $R = A_0/B_0$. Making use of the continued fractions, we obtain

$$\begin{aligned}R = \frac{e^{i2k_zd_0}}{R_{01}} &+ \cfrac{\left[1-(1/R_{01}^2)\right]e^{i2(k_{1z}+k_z)d_0}}{(1/R_{01})e^{i2k_{1z}d_0}} + \frac{e^{i2k_{1z}d_1}}{R_{12}} \\ &+ \cfrac{\left[1-(1/R_{12}^2)\right]e^{i2(k_{2z}+k_{1z})d_1}}{(1/R_{12})e^{i2k_{2z}d_1}} + \cdots + \frac{e^{i2k_{(n-1)z}d_{n-1}}}{R_{(n-1)n}} \\ &+ \cfrac{\left[1-(1/R_{(n-1)n}^2)\right]e^{i2(k_{nz}+k_{(n-1)z})d_{n-1}}}{(1/R_{(n-1)n})e^{i2k_{nz}d_{n-1}}} + R_{nt}e^{i2k_{nz}d_n}\end{aligned} \tag{24}$$

This is a closed-form solution for the reflection coefficient expressed in continued fractions. Such a solution is very easily programmed for numerical computation.

For TM plane waves reflected from the stratified medium, the principle of duality applies and gives rise to the answer for R identical to (24). The only difference is that (19) now becomes

$$p_{l(l+1)} = \frac{\epsilon_l k_{(l+1)z}}{\epsilon_{l+1}k_{lz}} \tag{25}$$

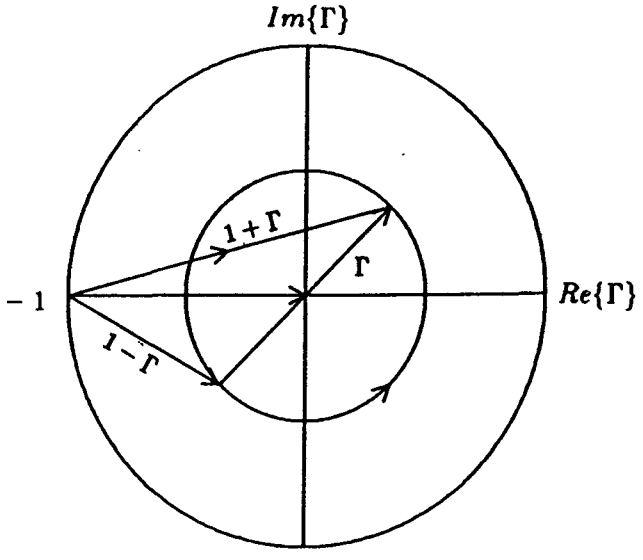


Figure 3.3.2 Complex Γ plane.

for TM waves. Thus, for the definition of the reflection coefficients in (20), we shall use (25) instead of (19).

For a stratified medium with any given number of layers (the number of layers is defined to be equal to t or equivalently equal to the number of boundaries), the reflection coefficient is derived from (24) by taking terms, starting from the last one, until the subscript l of $R_{l(l+1)}$ becomes zero. Consider, for instance, a two-layer medium, with $t = 2$ and $n = 1$. We find from (24), writing $k_{0z} = k_z$,

$$R = \frac{e^{i2k_z d_0}}{R_{01}} + \frac{\left[1 - (1/R_{01}^2)\right] e^{i2(k_{1z}+k_z)d_0}}{(1/R_{01})e^{i2k_{1z}d_0} + R_{12}e^{i2k_{1z}d_1}}$$

$$= \frac{R_{01} + R_{12}e^{i2k_{1z}(d_1-d_0)}}{1 + R_{01}R_{12}e^{i2k_{1z}(d_1-d_0)}} e^{i2k_z d_0} \tag{26}$$

Note that when $R_{01} = \pm 1$, the reflection coefficient in (24) will have magnitude unity, disregarding the composition of the stratified medium below $z = -d_0$. This should be the case, as $|R_{01}| = 1$ represents, for instance, a perfectly conducting coating.

From (7) – (9) we see that A_l/B_l is the ratio of the amplitude of the wave propagating in the positive $\hat{z}$ direction to that of the wave propagating in the negative $\hat{z}$ direction.

We define a space-dependent complex reflection coefficient $\Gamma_l(z)$ such that

$$\Gamma_l(x) = \frac{A_l}{B_l} e^{i2k_{lz}z}$$

On the complex $\Gamma_l(z)$ plane [Fig. 3.3.2], as the phase $\phi = 2k_{lz}z$ increases with z, $\Gamma_l(z)$ varies in a counterclockwise manner. If k_{lz} is complex, $\Gamma_l(z)$ decreases with increasing z.

We can define a wave impedance $Z_l(z)$ in the negative $\hat{z}$ direction:

$$Z_{lz}(z) = \frac{E_{ly}}{H_{lx}} = \frac{\omega\mu_l}{k_{lz}} \frac{1+\Gamma_l(z)}{1-\Gamma_l(z)}$$

which is complex. For a plane wave propagating in free space in the absence of any medium, the wave impedance in the direction of wave propagation is $\eta = \omega\mu_o/k = (\mu_o/\epsilon_o)^{1/2} \approx 377\,\Omega$.

With the definition of the complex wave impedance, the ratio of (15) to (16) gives $Z_{lz}(z = -d_l) = Z_{(l+1)z}(z = -d_l)$. Thus at each interface the wave impedance are continuous across the boundary.

On the complex Γ plane, $Z_{lz}(z)$ can be interpreted as the ratio of the two lengths as shown in Figure 3.3.2. The magnitude of $Z_{lz}(z)$ is maximum when Γ_l is real and positive. We define the dimensionless relative wave impedance as

$$z_l = \frac{Z_{lz}}{\omega\mu_l/k_{lz}} = \frac{1+\Gamma_l}{1-\Gamma_l}$$

For all possible complex values of $\Gamma_l(z)$, we can map the corresponding $z_l(z)$ values onto the complex $\Gamma_l(z)$ plane. The result is in the form of the Smith chart, which is frequently used in transmission line studies.

To illustrate the use of the wave impedance concept, consider a stratified medium composed of $2N+2$ isotropic dielectric layers (corresponding to $2N+2$ boundaries) with alternating high and low permittivites, ϵ_h and ϵ_l; regions $1, 3, 5, \ldots, 2N+1$ are high-permittivity layers, and regions $2, 4, 6, \ldots, 2N$ are low-permittivity layers. Region 0 has permittivity ϵ and permeability μ. The thickness of each layer is a quarter-wavelength inside the dielectric. The transmitted region is $2N+2 = t$ and has permittivity ϵ_t. Permeabilities for all layers are equal to μ [Fig. 3.3.3].

Consider a wave normally incident upon the stratified medium, $k_x = 0$, $k_{lz} = \omega\sqrt{\mu\epsilon_l}$ for all l. The wave impedance of region t,

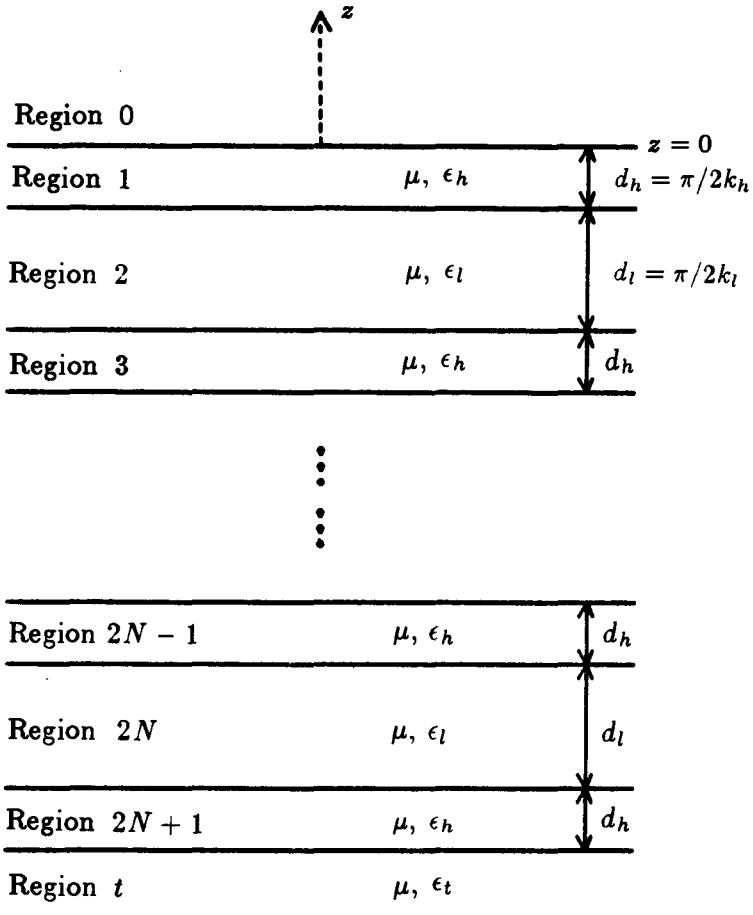


Figure 3.3.3 Layered medium with alternating high and low permittivities.

since there is no reflection, is $Z_t = (\mu/\epsilon_t)^{1/2}$. Because of the continuity of wave impedance across the boundary, the impedance across the interface separating regions $2N+1$ and t is $Z_{2N+1} = (\mu/\epsilon_t)^{1/2}$. The relative impedance is $z_{2N+1} = (\mu/\epsilon_t)^{1/2}/(\mu/\epsilon_h)^{1/2} = (\epsilon_h/\epsilon_t)^{1/2}$. Making use of the Smith chart concept, and noting the periodicity of the structure, we determine the wave impedance at $z = 0$:

$$Z_0 = \left(\frac{\epsilon_l}{\epsilon_h}\right)^{1/2} \left(\frac{\epsilon_l}{\epsilon_h}\right)^{N} \left(\frac{\mu}{\epsilon_h}\right)^{1/2}$$

The reflection coefficient R at $z = 0$ is found to be

$$R_0 = \frac{Z_0/(\mu/\epsilon)^{1/2} - 1}{Z_0/(\mu/\epsilon)^{1/2} + 1} = \frac{(\epsilon_t/\epsilon_h)^{1/2}(\epsilon_l/\epsilon_h)^N(\epsilon/\epsilon_h)^{1/2} - 1}{(\epsilon_t/\epsilon_h)^{1/2}(\epsilon_l/\epsilon_h)^N(\epsilon/\epsilon_h)^{1/2} + 1}$$

We observe that, for a high ϵ_h/ϵ_l ratio and for a larger number of layers, the reflection coefficient R_0 approaches the value -1, and the structure is highly reflective. Such structures are useful at optical frequencies since metallic reflectors are subject to corrosion and tarnishing problems.

b. Propagation Matrices and Transmission Coefficients

For a plane wave incident on a stratified medium, we have obtained the boundary conditions of continuity of tangential electric and magnetic fields at each interface $z = -d_l$, with the two equations (15)–(16) relating wave amplitudes in regions l and $l+1$:

$$A_{l+1}e^{-ik_{(l+1)z}d_l} + B_{l+1}e^{ik_{(l+1)z}d_l} = A_l e^{-ik_{lz}d_l} + B_l e^{ik_{lz}d_l} \tag{27}$$

$$A_{l+1}e^{-ik_{(l+1)z}d_l} - B_{l+1}e^{ik_{(l+1)z}d_l} = p_{(l+1)l}\left[A_l e^{-ik_{lz}d_l} - B_l e^{ik_{lz}d_l}\right] \tag{28}$$

where for TE waves

$$p_{(l+1)l} = \frac{\mu_{l+1}k_{lz}}{\mu_l k_{(l+1)z}} \tag{29}$$

and for TM waves

$$p_{(l+1)l} = \frac{\epsilon_{l+1}k_{lz}}{\epsilon_l k_{(l+1)z}} \tag{30}$$

Note that

$$p_{(l+1)l} = \frac{1}{p_{l(l+1)}} \tag{31}$$

Equations (29) and (30) follow from duality but bear in mind that A_l and B_l denote amplitudes of tangential electric fields for TE waves and denote amplitudes of tangential magnetic fields for TM waves. In the last section we determined the reflection coefficients $R = A_0/B_0$ from the $(2n+2)$ boundary conditions. We will now show that the transmission coefficient $T = B_t/B_0$ can be obtained by the use of propagation matrices.

We solve for A_{l+1} and B_{l+1} in terms of A_l and B_l from (27)–(28) and obtain

$$A_{l+1}e^{-ik_{(l+1)z}d_l} = \frac{1}{2}(1+p_{(l+1)l})\left(A_l e^{-ik_{lz}d_l} + R_{(l+1)l}B_l e^{ik_{lz}d_l}\right)$$

$$B_{l+1}e^{ik_{(l+1)z}d_l} = \frac{1}{2}(1+p_{(l+1)l})\left(R_{(l+1)l}A_l e^{-ik_{lz}d_l} + B_l e^{ik_{lz}d_l}\right)$$

Expressing in the form of matrix multiplication, we have

$$\begin{pmatrix} A_{l+1}e^{-ik_{(l+1)z}d_{l+1}} \\ B_{l+1}e^{ik_{(l+1)z}d_{l+1}} \end{pmatrix} = \overline{\overline{V}}_{(l+1)l} \cdot \begin{pmatrix} A_l e^{-ik_{lz}d_l} \\ B_l e^{ik_{lz}d_l} \end{pmatrix} \tag{32}$$

where

$$\overline{\overline{V}}_{(l+1)l} = \frac{1}{2}[1+p_{(l+1)l}] \begin{pmatrix} e^{-ik_{(l+1)z}(d_{l+1}-d_l)} & R_{(l+1)l}e^{-ik_{(l+1)z}(d_{l+1}-d_l)} \\ R_{(l+1)l}e^{ik_{(l+1)z}(d_{l+1}-d_l)} & e^{ik_{(l+1)z}(d_{l+1}-d_l)} \end{pmatrix} \tag{33}$$

is called the forward-propagating matrix. In (33),

$$R_{(l+1)l} = \frac{1-p_{(l+1)l}}{1+p_{(l+1)l}} = -R_{l(l+1)}$$

is the reflection coefficient at the boundary separating regions $l+1$ and l, and the first subscript denotes the region with the incident wave. It is to be noted for the forward-propagating matrix between layers n and $t = n+1$,

$$\begin{pmatrix} 0 \\ T \end{pmatrix} = \overline{\overline{V}}_{tn} \cdot \begin{pmatrix} A_n e^{-ik_{nz}d_n} \\ B_n e^{ik_{nz}d_n} \end{pmatrix}$$

with

$$\overline{\overline{V}}_{tn} = \frac{1}{2}(1+p_{tn}) \begin{pmatrix} e^{ik_{tz}d_n} & R_{tn}e^{ik_{tz}d_n} \\ R_{tn}e^{-ik_{tz}d_n} & e^{-ik_{tz}d_n} \end{pmatrix}$$

By the same token, we may express A_l and B_l in terms of A_{l+1} and B_{l+1} by using (17) – (18) and define a backward-propagating matrix.

The propagation matrices can be used to determine wave amplitudes in any region in terms of those in any other region. For $m > l$, we make use of the forward propagation matrix to obtain

$$\begin{pmatrix} A_m e^{-ik_{mz}d_m} \\ B_m e^{ik_{mz}d_m} \end{pmatrix} = \overline{\overline{V}}_{m(m-1)} \cdot \overline{\overline{V}}_{(m-1)(m-2)} \cdots \overline{\overline{V}}_{(l+1)l} \cdot \begin{pmatrix} A_l e^{-ik_{lz}d_l} \\ B_l e^{ik_{lz}d_l} \end{pmatrix}$$

Similarly, backward-propagating matrices can be used to express wave amplitudes in any region j in terms of those in region l for $l > j$.

In particular, the transmission coefficient $T = B_t/B_0$ for a stratified medium with $t = n+1$ layers can be calculated by the multiplication of $n+1$ propagation matrices. Using the forward-propagating matrices, we have

$$\begin{pmatrix} 0 \\ T \end{pmatrix} = \overline{\overline{V}}_{t0} \cdot \begin{pmatrix} Re^{-ik_z d_0} \\ e^{ik_z d_0} \end{pmatrix}$$

where

$$\overline{\overline{V}}_{t0} = \overline{\overline{V}}_{tn} \cdot \overline{\overline{V}}_{n(n-1)} \cdots \overline{\overline{V}}_{10}$$

includes all information about the stratified medium. Once $\overline{\overline{V}}_{t0}$ is known, both the reflection and transmission coefficients can be calculated from its matrix elements.

As a first example, we calculate the transmission coefficient for a one-layer (half-space) medium. From (33) and by letting $d_0 = 0$, we find

$$\begin{pmatrix} 0 \\ T \end{pmatrix} = \frac{1}{2}(1+p_{t0}) \begin{pmatrix} 1 & R_{t0} \\ R_{t0} & 1 \end{pmatrix} \begin{pmatrix} R_{0t} \\ 1 \end{pmatrix}$$

which gives

$$T = \frac{1}{2}(1+p_{t0})(1-R_{0t}^2) = \frac{2}{1+p_{0t}} \tag{34}$$

where we made use of the fact that $P_{t0} = 1/P_{0t}$ and $R_{t0} = -R_{0t}$. The implications of these reflection and transmission coefficients for both TE and TM waves have been discussed in the last section.

For a two-layer medium with $d_0 = 0$, we obtain the transmission coefficient from

$$\begin{pmatrix} 0 \\ T \end{pmatrix} = \frac{1}{4}(1+p_{t1})(1+p_{10}) \begin{pmatrix} e^{ik_{tz}d_1} & R_{t1}e^{ik_{tz}d_1} \\ R_{t1}e^{-ik_{tz}d_1} & e^{-ik_{tz}d_1} \end{pmatrix} \begin{pmatrix} e^{-ik_{1z}d_1} & R_{10}e^{-ik_{1z}d_1} \\ R_{10}e^{ik_{1z}d_1} & e^{ik_{1z}d_1} \end{pmatrix} \begin{pmatrix} R \\ 1 \end{pmatrix}$$

which yields

$$\begin{aligned} T &= \frac{1}{4}(1+p_{10})(1+p_{1t}) \frac{(1-R_{01}^2)(1-R_{1t}^2)}{1+R_{01}R_{1t}e^{i2k_{1z}d_1}} e^{i(k_{1z}-k_{tz})d_1} \\ &= \frac{4e^{i(k_{1z}-k_{tz})d_1}}{(1+p_{01})(1+p_{1t})(1+R_{01}R_{1t}e^{i2k_{1z}d_1})} \end{aligned} \tag{35}$$

The reflectivity and the transmissivity for the two-layer medium, in view of (3.2.23) and (3.2.24), are found to be

$$r = |R|^2$$

$$t = p_{0t}\,|T|^2$$

It is straightforward to prove that $r + t = 1$ with R and T given by (26) and (35).

As another example we use the forward-propagation matrix formalism to determine the transmission and the reflection coefficients for a periodic medium made of $2N+2$ isotropic dielectric layers with alternating high and low permittivities ϵ_h and ϵ_l [Fig. 3.3.3]. The thickness of each layer is a quarter-wavelength inside the dielectric. The transmitted region is $t = 2N+2$ and has permittivity ϵ_t. Consider normal incidence, $k_x = 0$. The reflection coefficient R has been calculated with the wave impedance approach. Using the forward-propagation matrix formalism, we have

$$\begin{pmatrix} 0 \\ T \end{pmatrix} = \overline{\overline{V}}_{th} \cdot \left(\overline{\overline{V}}_{hl} \cdot \overline{\overline{V}}_{lh}\right)^N \cdot \overline{\overline{V}}_{k0} \cdot \begin{pmatrix} R \\ 1 \end{pmatrix}$$

Since in region $m+1$, $k_{(m+1)z} = k_{m+1}$ and $(d_{m+1} - d_m)$ is a quarter-wavelength thick, we have $k_{(m+1)z}(d_{m+1} - d_m) = \pi/2$. Note also the fact that $\mu_{m+1}\, k_{mz}/\mu_m\, k_{(m+1)z} = (\epsilon_m/\epsilon_{m+1})^{1/2}$. The forward-propagation matrices become

$$\overline{\overline{V}}_{ho} = -\frac{i}{2}\begin{pmatrix} 1+\sqrt{\epsilon/\epsilon_h} & 1-\sqrt{\epsilon/\epsilon_h} \\ -1+\sqrt{\epsilon/\epsilon_h} & -1-\sqrt{\epsilon/\epsilon_h} \end{pmatrix}$$

$$\overline{\overline{V}}_{hl} \cdot \overline{\overline{V}}_{lh} = -\frac{1}{2}\begin{pmatrix} \sqrt{\epsilon_l/\epsilon_h}+\sqrt{\epsilon_h/\epsilon_l} & \sqrt{\epsilon_l/\epsilon_h}-\sqrt{\epsilon_h/\epsilon_l} \\ \sqrt{\epsilon_l/\epsilon_h}-\sqrt{\epsilon_h/\epsilon_l} & \sqrt{\epsilon_l/\epsilon_h}+\sqrt{\epsilon_h/\epsilon_l} \end{pmatrix}$$

$$\overline{\overline{V}}_{th} = \frac{1}{2}\begin{pmatrix} (1+\sqrt{\epsilon_h/\epsilon_t})e^{ik_t d} & (1-\sqrt{\epsilon_h/\epsilon_t})e^{ik_t d} \\ (1-\sqrt{\epsilon_h/\epsilon_t})e^{-ik_t d} & (1+\sqrt{\epsilon_h/\epsilon_t})e^{-ik_t d} \end{pmatrix}$$

where d is the total thickness of the periodic medium. The term $(\overline{\overline{V}}_{hl} \cdot \overline{\overline{V}}_{lh})^N$ can be calculated by making use of the matrix identity

$$\begin{pmatrix} a+b & a-b \\ a-b & a+b \end{pmatrix}^N = 2^{N-1}\begin{pmatrix} a^N+b^N & a^N-b^N \\ a^N-b^N & a^N+b^N \end{pmatrix}$$

It follows that

$$R = \frac{(\epsilon_l/\epsilon_h)^N - \sqrt{\epsilon_h^2/\epsilon\epsilon_t}}{(\epsilon_l/\epsilon_h)^N + \sqrt{\epsilon_h^2/\epsilon\epsilon_t}}$$

and

$$T = \frac{2i(-1)^N(\epsilon/\epsilon_t)^{1/2}e^{-ik_t d}}{\sqrt{\epsilon/\epsilon_h}\,(\epsilon_l/\epsilon_h)^{N/2} + \sqrt{\epsilon_h/\epsilon_t}\,(\epsilon_h/\epsilon_l)^{N/2}}$$

In view of (3.2.23) – (3.2.24) we find that the reflectivity $r = |R|^2$, and the transmissivity $t = (\epsilon_t/\epsilon)^{1/2}|T|^2$. Again it can be shown that $r + t = 1$. Note that, although both TE and TM waves become TEM at normal incidence, we must use $p_{0t} = k_{tz}/k_z = (\epsilon_t/\epsilon)^{1/2}$ because here R and T are amplitude reflection and transmission coefficients for electric field vectors.

3.4 Guidance by Conducting Parallel Plates

Consider the guidance of electromagnetic waves by a pair of perfectly conducting plates at $x = 0$ and $x = -d$ [Fig. 3.4.1]. The medium between the plates is homogeneous and isotropic. The width of the waveguide along y is w and we assume $w \gg d$, such that fringing fields can be neglected, and we have $\partial/\partial y = 0$. The Maxwell equations can be decomposed into transverse electric (TE) and transverse magnetic (TM) components. We have

$$H_x = -\frac{1}{i\omega\mu}\frac{\partial}{\partial z}E_y \tag{1a}$$

$$H_z = \frac{1}{i\omega\mu}\frac{\partial}{\partial x}E_y \tag{1b}$$

$$\left(\frac{\partial^2}{\partial z^2} + \frac{\partial^2}{\partial x^2} + \omega^2\mu\epsilon\right)E_y = 0 \tag{1c}$$

for TE waves and

$$E_x = \frac{1}{i\omega\epsilon}\frac{\partial}{\partial z}H_y \tag{2a}$$

$$E_z = -\frac{1}{i\omega\epsilon}\frac{\partial}{\partial x}H_y \tag{2b}$$

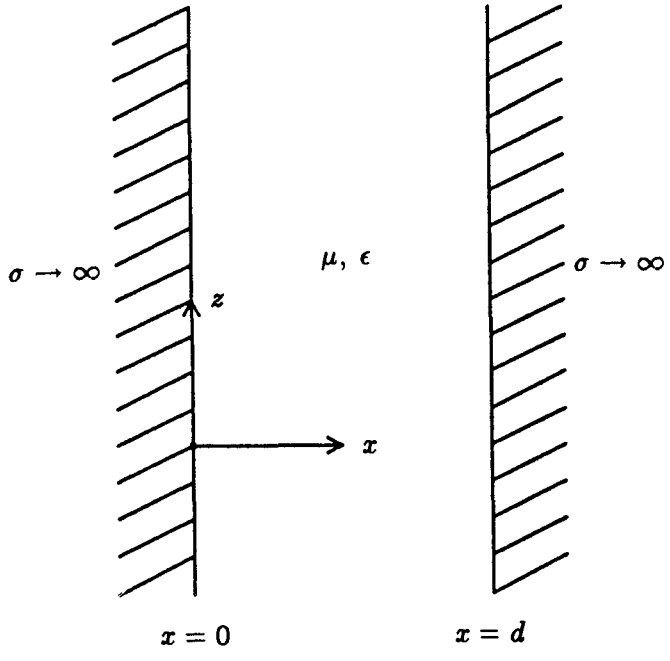


Figure 3.4.1 Parallel-plate waveguide.

$$\left(\frac{\partial^2}{\partial z^2}+\frac{\partial^2}{\partial x^2}+\omega^2\mu\epsilon\right)H_y=0 \tag{2c}$$

for TM waves. Equations (1*a*)–(1*c*) are duals of (2*a*)–(2*c*). However, the boundary conditions at the parallel plates require a zero tangential electric field for TE waves and also for TM waves. The boundary conditions for TE and TM waves are not duals.

The wave is guided along $\pm\hat{z}$ directions. For waves propagating along the $+\hat{z}$ direction, the z dependence of all field vectors is $e^{ik_z z}$, where we call k_z the propagation constant. For TE waves, the solution for E_y, satisfying the boundary conditions, is

$$E_y = E_0 \sin k_x x\, e^{ik_z z} \tag{3}$$

where $E_y = 0$ at the boundary $x = 0$. In order to have $E_y = 0$ at the boundary $x = d$, we find

$$k_x = \frac{m\pi}{d} \tag{4}$$

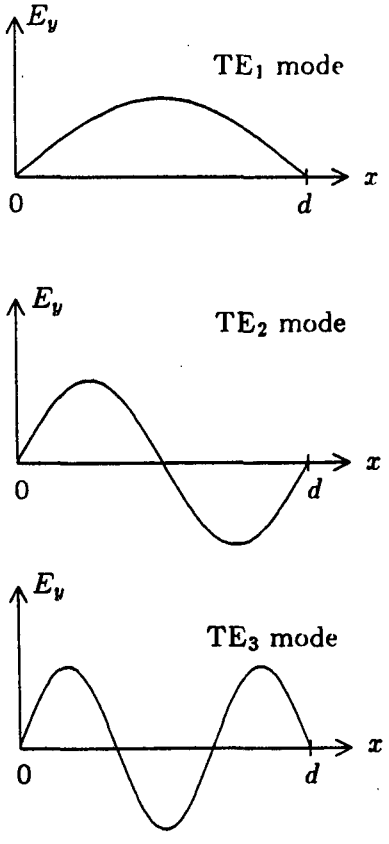


Figure 3.4.2 Field amplitudes for TE_1, TE_2, and TE_3 modes.

where m is any integer. Equation (4) is called the *guidance condition* which is determined from the boundary conditions. Substituting (3) in (1c) we find the dispersion relation

$$k_z^2 + k_x^2 = \omega^2 \mu \epsilon = k^2 \tag{5}$$

We call the TE wave corresponding to each integer m the TE_m mode. There is no TE_0 mode because $E_y = 0$ for $m = 0$. The field patterns for E_y are plotted in Figure 3.4.2 for $m = 1, 2,$ and 3.

Substituting the guidance condition (4) in the dispersion relation (5) we obtain

$$k_z^2 + \left(\frac{m\pi}{d}\right)^2 = k^2 \tag{6}$$

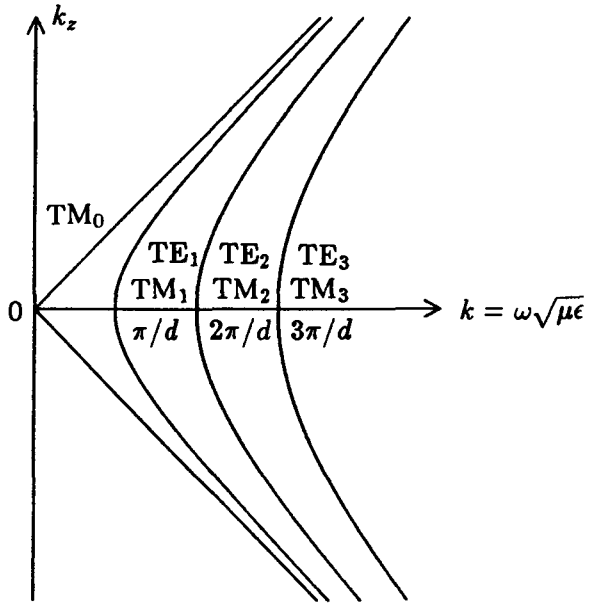


Figure 3.4.3 $k_z - k$ **diagram for guided modes.**

This equation describes a family of hyperbolas for different values of m. In Figure 3.4.3 we plot the $k_z - k$ diagram. We see that for the mth mode k_z will be imaginary if $k < m\pi/d$. The wave then becomes evanescent and attenuates exponentially in the $\hat{z}$ direction. The wavenumber at which $k_z = 0$ is called the cutoff wavenumber

$$k_{cm} = m\pi/d \tag{7}$$

The corresponding cutoff angular frequency is $\omega_{cm} = k_{cm}/(\mu\epsilon)^{1/2} = m\pi/d(\mu\epsilon)^{1/2}$ and the corresponding cutoff wavelength is $\lambda_{cm} = 2d/m$. In order for the mth order TE mode to propagate, k must be larger than k_{cm}. Notice that if the TM_m mode is propagating, then all TE_l modes with $l < m$ can also propagate. Thus for a given wavenumber k such that $m\pi/d < k < (m+1)\pi/d$, there will be m TE modes admissible inside the waveguide. The lowest order TE mode is TE_1 whose $k_{c1} = \pi/d$. Thus for $k < \pi/d$, no TE mode can be excited. The single TE_1 mode operation range inside the guide is $\pi/d < k < 2\pi/d$.

Below the cutoff frequency for the mth mode, $k_z = i\alpha_z$ and (6)

give

$$\alpha_z^2 + k^2 = \left(\frac{m\pi}{d}\right)^2$$

Above the cutoff frequency, the phase velocity v_p along the $\hat{z}$ direction is

$$v_p = \frac{\omega}{k_z} = \frac{1}{\sqrt{\mu\epsilon}}\left[1 - \left(\frac{k_{cm}}{k}\right)^2\right]^{-1/2}$$

As the frequency goes to infinity, v_p approaches the velocity of light in the medium. For finite ω, the phase velocity is always larger than $(\mu\epsilon)^{-1/2}$. This is because k_z is always smaller than k. The fact that the phase velocity depends on frequency makes the waveguide a dispersive transmission system. The group velocity of propagation is

$$v_g = \left(\frac{\partial k_z}{\partial \omega}\right)^{-1} = \frac{1}{\sqrt{\mu\epsilon}}\left[1 - \left(\frac{k_{cm}}{k}\right)^2\right]^{1/2}$$

Thus $v_p v_g = 1/\mu\epsilon$. The group velocity is always smaller than the velocity of light inside the medium. In fact,

$$v_g = \frac{1}{\sqrt{\mu\epsilon}}\frac{k_z}{k}$$

which is equal to the projection of the velocity of wave in the medium into the $\hat{z}$ direction of propagation.

The field solution (3) has a simple physical interpretation. Writing in terms of exponential functions instead of sine, we have

$$E_y = \frac{E_0}{2i}\left[e^{ik_x x + ik_z z} - e^{-ik_x x + ik_z z}\right] \tag{8}$$

which represents a superposition of two plane waves. We may view the guided wave as resulting from a plane wave bouncing back and forth between the two plates and propagating in the $\hat{z}$ direction with the propagation constant k_z [Fig. 3.4.4].

The magnetic field vector is obtained from (3) by using Faraday's law

$$\begin{aligned}\overline{H} &= \frac{1}{i\omega\mu}\left(-\hat{x}\frac{\partial}{\partial z}E_y + \hat{z}\frac{\partial}{\partial x}E_y\right) \\ &= \frac{E_0}{i\omega\mu}\left(-\hat{x}ik_z \sin k_x x + \hat{z}k_x \cos k_x x\right) e^{ik_z z}\end{aligned}$$

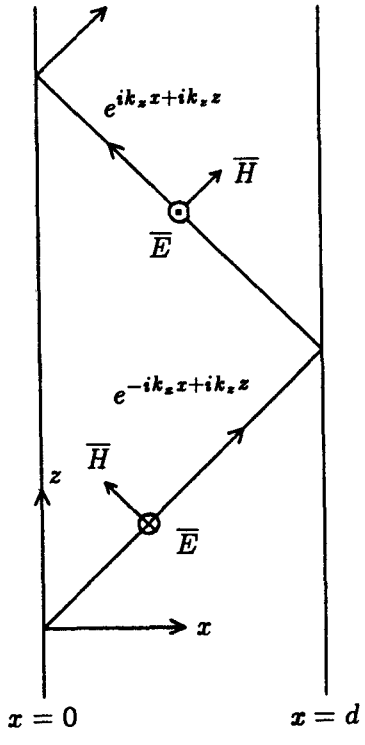


Figure 3.4.4 Plane wave interpretation of guided modes.

Casting in the plane wave representation, we find

$$\overline{H} = \frac{E_0}{i2\omega\mu}\Big[(-\hat{x}k_z + \hat{z}k_x)e^{ik_x x+ik_z z} + (\hat{x}k_z + \hat{z}k_x)e^{-ik_x x+ik_z z}\Big] \tag{9}$$

We see that the magnetic field vector is perpendicular to both the electric field vector and the wave vector for the two bouncing plane waves.

The guidance condition (4) states that in the $\hat{x}$ direction the bouncing waves must interfere constructively with $2k_x d = 2m\pi$ in order for the wave to be guided. Since there is only a set of discrete k_x values admissible inside the guide, the corresponding k_z values are determined from the dispersion relation (6) and are shown in Figure 3.4.5.

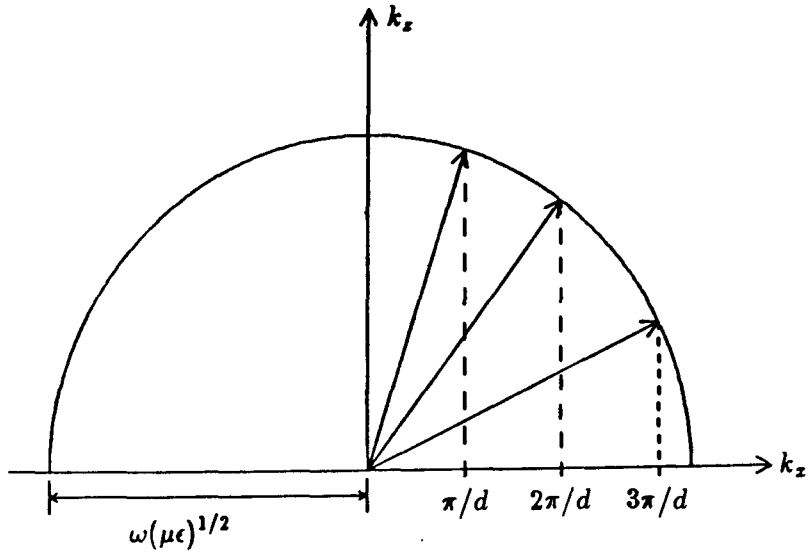


Figure 3.4.5 Interpretation of the guidance condition.

Below cutoff, both the phase and group velocities are zero. We can see from the field solutions that there is no time-average Poynting power flow in the waveguide. From (3) we can find the associated magnetic fields from Faraday's law

$$\overline{H} = -E_0 \left(\hat{x} \frac{k_z}{\omega\mu} \sin \frac{m\pi x}{d} + \hat{z} i \frac{m\pi}{\omega\mu d} \cos \frac{m\pi x}{d} \right) e^{ik_z z} \tag{10}$$

The complex Poynting power along $\hat{z}$ is

$$S_z = \frac{k_z^*}{\omega\mu} |E_0|^2 \sin^2 \frac{m\pi x}{d} e^{i(k_z - k_z^*)z}$$

When k_z is imaginary, $k_z = i\alpha_z$, we see that the time-average Poynting power $< S_z >= \frac{1}{2} Re(S_z) = 0$. The time-average Poynting power $< S_x >$ in the transverse direction is always zero because the complex conjugate of the z component of $\overline{H}$ in (10) multiplied by E_y is purely imaginary.

For TM waves satisfying the boundary conditions, we find that the solution takes the form

$$\overline{H} = \hat{y} H_0 \cos k_x x \, e^{ik_z z} \tag{11a}$$

$$\overline{E} = H_0 \frac{1}{\omega\epsilon} \left[\hat{x} k_z \cos k_x x - \hat{z} i k_x \sin k_x x \right] e^{ik_z z} \tag{11b}$$

The boundary condition of vanishing E_z at $x = 0$ and $x = d$ leads to the guidance condition

$$k_x = \frac{m\pi}{d} \tag{12}$$

Substituting the solution for H_y in the wave equation (2) yields the dispersion relation

$$k_z^2 + k_x^2 = \omega^2\mu\epsilon = k^2$$

The dispersion relation after the substitution of the guidance condition is identical to (6) and is plotted in Figure 3.4.5. Both the guidance condition (11) and the dispersion relation (12) are identical to those for TE modes except for one very important difference. When $m = 0$, the field no longer vanishes as in the TE case. Therefore we can now have the TM_0 mode which is also called the TEM mode. The cutoff frequency for the TEM mode is zero. Thus the TEM mode in a parallel-plate waveguide is termed the fundamental or dominant mode with a single-mode operation range that extends from $k = 0$ to $k = \pi/d$.

Field solutions for the TM_0 mode follows from (11) when we set $k_x = 0$ and $k_z = k$. We have

$$\overline{H} = \hat{y}H_0\, e^{ikz} \tag{13a}$$

$$\overline{E} = \hat{x}\eta H_0\, e^{ikz} \tag{13b}$$

where $\eta = \sqrt{\mu/\epsilon}$ is the characteristic impedance of the medium. The electric field is seen to be perpendicular to the plates and the magnetic field parallel to the plates.

a. Transmission Line Theory

The transmission line theory is best illustrated with the TM_0 mode in parallel-plate waveguides. To accomodate the cases when the waveguide may be terminated at one end, we write the solution (13) in a more general form

$$\overline{E} = \hat{x}E_x(z) \tag{14a}$$

$$\overline{H} = \hat{y}H_y(z) = \hat{y}\frac{1}{\eta}E_x(z) \tag{14b}$$

Assume that the width of the parallel-plate waveguide in the $\hat{y}$ direction is w and assume that fringing fields in the $\hat{y}$ direction may

be neglected so that $\partial/\partial y = 0$. We can define, at a constant z, the voltage $V(z)$

$$V(z) = V(x=0) - V(x=d) = \int_0^d dx\, E_x = E_x d \tag{15a}$$

and the current $I(z)$

$$I(z) = \int_0^w dy\, \hat{z} \cdot (\hat{x} \times \overline{H}) = H_y w \tag{15b}$$

The Maxwell equation

$$\nabla \times \overline{E} = i\omega\mu\overline{H}$$

then yields

$$\frac{d}{dz}V(z) = i\omega L I(z) \tag{16a}$$

where $L = \mu d/w$ (h/m) is the inductance per unit length for the parallel-plate waveguide. Similarly, the Maxwell equation

$$\nabla \times \overline{H} = -i\omega\epsilon\overline{E}$$

yields

$$\frac{d}{dz}I(z) = i\omega C V(z) \tag{16b}$$

where $C = \epsilon w/d$ (f/m) is the capacitance per unit length for the parallel-plate waveguide.

Equations (16a) and (16b) are referred to as the transmission line equations. Notice that for guided modes other than TM_0, (16) cannot be derived as there will be x dependence. Eliminating $I(z)$ from (16a) and (16b), we find the homogeneous Helmholtz equation

$$\left(\frac{d^2}{dz^2} + \omega^2 LC\right)V(z) = 0$$

For a wave travelling in the positive $\hat{z}$ direction, we write

$$V(z) = V_0 e^{ikz}$$

The corresponding current is determined from (16a)

$$I(z) = \frac{k}{\omega L}V(z) = \frac{1}{Z_0}V(z)$$

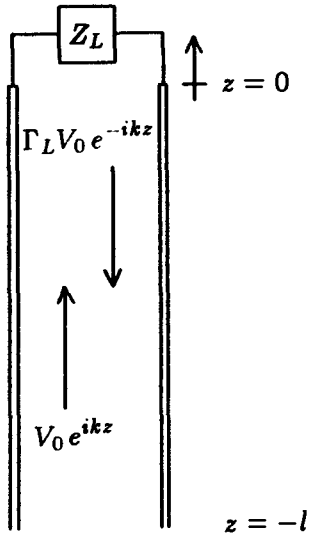


Figure 3.4.6 Transmission line with termination impedance Z_L.

where

$$Z_0 = \sqrt{\frac{L}{C}} \tag{17}$$

is called the characteristic impedance of the transmission line.

When the transmission line is terminated with an impedance Z_L [Fig. 3.4.6], there will be negatively travelling waves as well as positively travelling waves. We let Z_L be located at $z = 0$ and write

$$V(z) = V_0(e^{ikz} + \Gamma_L e^{-ikz}) \tag{18a}$$

where Γ_L is the reflection coefficient to be determined by the load impedance Z_L.

From (16a), we find

$$I(z) = \frac{V_0}{Z_0}(e^{ikz} - \Gamma_L e^{-ikz}) \tag{18b}$$

A generalized impedance $Z(z)$ is defined as, in view of (18),

$$Z(z) \equiv \frac{V(z)}{I(z)} = Z_0 \frac{1 + \Gamma_L e^{-i2kz}}{1 - \Gamma_L e^{-i2kz}} \tag{19}$$

At $z = 0$, we have

$$Z_L = Z(z=0) = Z_0 \frac{1+\Gamma_L}{1-\Gamma_L}$$

which yields the solution for Γ_L,

$$\Gamma_L = \frac{Z_{Ln}-1}{Z_{Ln}+1} \tag{20}$$

where

$$Z_{Ln} \equiv \frac{Z_L}{Z_0}$$

is the normalized load impedance. It is easily seen from (20) that (i) for a short circuit termination with $Z_L = 0$, we find $\Gamma_L = -1$; (ii) for an open circuit termination with $Z_L \to \infty$, we have $\Gamma_L = 1$; and (iii) for a matched termination with $Z_L = Z_0$, there will be no reflection at all and $\Gamma_L = 0$.

A generalized reflection coefficient $\Gamma(z)$ is defined as

$$\Gamma(z) \equiv \Gamma_L e^{-i2kz} \tag{21}$$

From (19) we can define a normalized impedance

$$Z_n(z) = \frac{Z(z)}{Z_0} = \frac{1+\Gamma(z)}{1-\Gamma(z)} \tag{22}$$

A Smith chart can be constructed from (22) for $\Gamma(z)$ taking all complex values with magnitude $|\Gamma(z)| \leq 1$. Within the unit circle on the complex Γ plane the loci on the Smith chart represent constant resistive and reactive values.

In terms of the generalized reflection coefficient $\Gamma(z)$, (18a) gives

$$V(z) = V_0 e^{ikz}[1+\Gamma(z)] \tag{23}$$

from which we can construct a voltage standing wave pattern as shown in Figure 3.4.7. The variation of the magnitude $|1+\Gamma(z)|$ is shown in Figure 3.4.7c. At $z = 0$, $\Gamma(z=0) = \Gamma_L$ which is a complex number determined from (20) by the normalized load impedance $Z_{Ln} = Z_L/Z_0$. When z becomes more negative as we move away from the load in

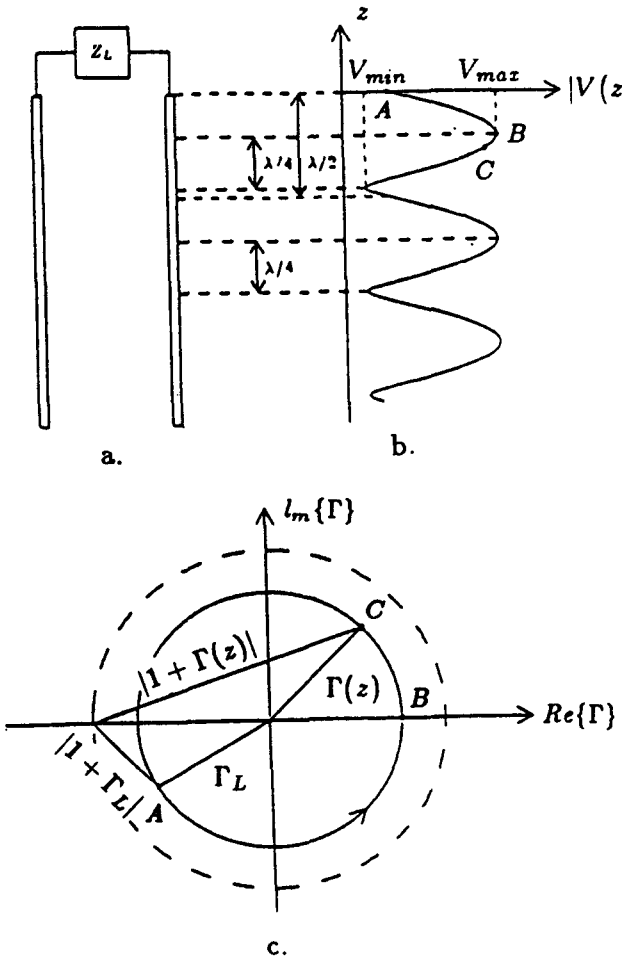


Figure 3.4.7 Construction of the voltage standing wave pattern.

the negative $\hat{z}$ direction, the locus of the generalized reflection coefficient in (21) rotates in a counter-clockwise direction on the complex Γ plane. The corresponding magnitude of $|V(z)|$ as plotted in Figure 3.4.7b is proportional to $|1+\Gamma(z)|$. The points A, B, and C in Figure 3.4.7 correspond to the same points in Figure 3.4.7b. As $k_z = \pi$, the voltage standing wave pattern repeats itself periodically. The maximum magnitude V_{max} occurs as $\Gamma(z)$ becomes positive real and V_{min} occurs as $\Gamma(z)$ becomes negative real. The voltage standing wave ratio

(VSWR) is defined as

$$\text{VSWR} = \frac{V_{max}}{V_{min}} = \frac{1 + |\Gamma(z)|}{1 - |\Gamma(z)|}$$

It is seen that for a matched load where there is no reflection, $\Gamma_L = 0$ and $\text{VSWR} = 1$. In general, VSWR takes values between unity and infinity.

b. Excitation of Modes in Parallel-Plate Waveguides

We now consider the excitation of TE and TM modes inside a parallel-plate waveguide. The mode amplitudes of the guided waves are determined by the sources of external excitation. We assume that the source is a current sheet located at $z = 0$. This current sheet generates propagating as well as evanescent guided modes in both positive and negative $\hat{z}$ directions. The boundary conditions at $z = 0$ require that (i) E_x and E_y be continuous, (ii) the discontinuity in H_x be equal to a current sheet flowing in the $\hat{y}$ direction, and (iii) the discontinuity in H_y be equal to a current sheet flowing in the $\hat{x}$ direction.

As an example, consider a current sheet in the $\hat{y}$ direction at $z = 0$ and varying in amplitude in x,

$$\overline{J}_s = \hat{y} J_s(x)$$

This current sheet can be visualized as composed of closely aligned wires with each wire excited by a different current source. The boundary conditions then require that

$$H_x \,|_{z=0_+} - H_x \,|_{z=0_-} = J_s(x)$$

and that all other tangential field components be continuous at $z = 0$.

As an example, we consider the following surface current. We write

$$\overline{J} = \hat{y} I_0 \delta(x - a)$$

which represents a line source located at $x = a$ and flowing in the $\hat{y}$ direction with the dimension ampere per meter. According to the boundary conditions, only TE waves will be excited. One can actually assume some amplitudes for the TM modes and find out from the

boundary conditions that they are all zero. We write the TE solutions as a superposition of all TE modes

$$E_y = \begin{cases} \sum_{m=1}^{\infty} E_m \sin \dfrac{m\pi x}{d} e^{ik_z z} & z \geq 0 \\ \sum_{m=1}^{\infty} E_m \sin \dfrac{m\pi x}{d} e^{-ik_z z} & z \leq 0 \end{cases} \tag{24}$$

We see that the boundary condition of E_y continuous at $z = 0$ has been satisfied. The amplitudes E_m in regions $z < 0$ and $z > 0$ are equal as a consequence of symmetry. The x components of the magnetic fields are

$$H_x = \begin{cases} \sum_{m=1}^{\infty} \dfrac{-k_z}{\omega\mu} E_m \sin \dfrac{m\pi x}{d} e^{ik_z z} & z \geq 0 \\ \sum_{m=1}^{\infty} \dfrac{k_z}{\omega\mu} E_m \sin \dfrac{m\pi x}{d} e^{-ik_z z} & z \leq 0 \end{cases} \tag{25}$$

At $z = 0$ the boundary condition gives

$$I_0 \delta(x - a) = \sum_{m=1}^{\infty} -\frac{2k_z}{\omega\mu} E_m \sin \frac{m\pi x}{d} \tag{26}$$

Using orthogonality properties of sinusoidal functions, we multiply both sides by $\sin(m\pi x/d)$ and integrate from 0 to d. The mode amplitude E_m is determined as

$$E_m = -\frac{\omega\mu}{k_z d} I_0 \sin \frac{m\pi a}{d} \tag{27}$$

For the TE_1 mode, E_1 is maximum when $a = d/2$. This is because E_y is also maximum at $x = d/2$ and the coupling of source energy into the TE_1 mode is the largest.

The time-average Poynting's power propagating along the waveguide in the $\hat{z}$ direction is given by

$$\begin{aligned} P &= \frac{1}{2} Re \left\{ \int_0^d dx \int_0^w dy \left(\sum_{m=1}^{\infty} E_m \sin \frac{m\pi x}{d} e^{ik_z z} \right) \right. \\ &\qquad \left. \cdot \left(\sum_{m=1}^{\infty} \frac{k_z}{\omega\mu} E_m \sin \frac{m\pi x}{d} e^{ik_z z} \right)^* \right\} \\ &= \frac{1}{2} Re \left\{ \frac{wd}{2} \sum_{m=1}^{\infty} |E_m|^2 \left(\frac{k_z^*}{\omega\mu} \right) e^{i(k_z - k_z^*)z} \right\} \end{aligned} \tag{28}$$

If the wavelength λ is such that $\lambda_{C(N+1)} < \lambda < \lambda_{CN} = 2d/N$, then k_z is real for $m \leq N$ and imaginary for $m > N$. Equation (28) gives

$$P = \sum_{m=1}^{N} \frac{wd}{4\eta} |E_m|^2 \sqrt{1 - \left(\frac{k_{cm}}{k}\right)^2} \tag{29}$$

where $\eta = \sqrt{\mu/\epsilon}$ is the intrinsic impedance of the medium inside the waveguide. The total power inside the guide is a summation of those of all the propagating modes with real k_z. It is important to observe that there is no coupling among the various modes; each individual mode carries its own power.

c. Attenuation of Guided Waves Due to Wall Loss

The flow of power inside the waveguide will not be attenuated if the parallel plates are indeed perfect conductors and the medium is perfectly lossless. Let us now investigate wave attenuation when the conductivity of the plates is large but finite. A perturbation approach will be used.

Due to wall loss, guided waves carrying total time-average power P_f will decrease as a function of z. Assume that the amplitudes of the fields decay exponentially with an attenuation constant α, where α is a small number. When the walls are perfectly conducting, $\alpha = 0$. The total propagating power P_f will decay exponentially with 2α such that $P_f \sim e^{-2\alpha z}$. By the power conservation principle, the decrease of P_f as a function of distance must be equal to the power dissipated per unit length P_d. We have

$$P_d = -\frac{d}{dz} P_f = 2\alpha P_f$$

The attenuation constant α is found to be

$$\alpha = \frac{P_d}{2P_f} \tag{30}$$

The objective is to calculate α by a perturbation approach.

The starting point of the perturbation process is the exact solution for perfectly conducting waveguides. We calculate the time-average power flow in the $\hat{z}$ direction P_f by using the unperturbed solutions for the fields

$$P_f = \frac{1}{2} Re \left\{ \iint dx\, dy\, \hat{z} \cdot (\overline{E} \times \overline{H}^*) \right\} \tag{31}$$

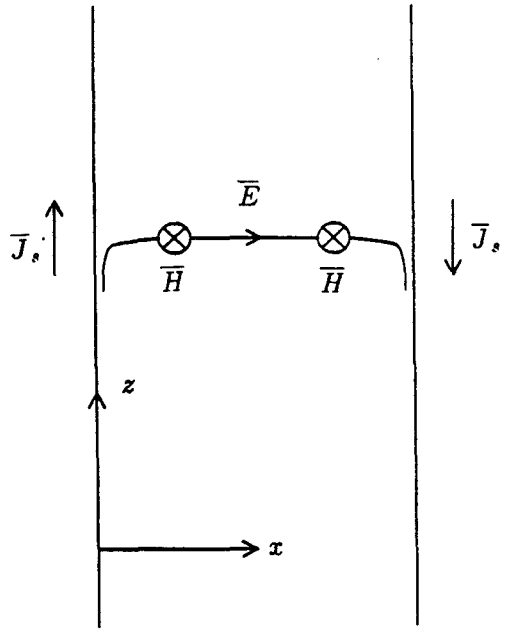


Figure 3.4.8 Finitely conducting walls.

The integration is carried over the area perpendicular to the direction of propagation.

To estimate the dissipated power per unit length P_d, we first investigate the origin of the dissipation due to imperfect wall conductivity. The surface currents on the plate surfaces are

$$\overline{J}_s = \hat{n} \times \overline{H}_W \quad \text{amp/m} \tag{32}$$

where $\overline{H}_W$ is the magnetic field at the walls for which $x = 0$ and d. When the walls are perfectly conducting, the tangential electric fields are zero and the surface currents flow without the support of any electric field. When the walls are not perfectly conducting, there is a small tangential component of the electric field $\overline{E}_W$ to support the surface currents [Fig. 3.4.8]. The time-average dissipated power per unit area into the conductors is, from Poynting's theorem, $< \overline{E}_W \cdot \overline{J}_s^* >$ watts/m^2. We shall calculate the dissipated power in terms of the

unperturbed value $\overline{H}_w$. The electric and magnetic fields at the wall surface are related to each other by the intrinsic impedance of the conductor which is assumed to have a large conductivity σ such that

$$\overline{E}_w = \sqrt{\mu/\epsilon_w}\,\hat{n} \times \overline{H}_w \approx \sqrt{\omega\mu/i\sigma}\,\hat{n} \times \overline{H}_w$$

where we assume the same permeability μ for the conductor as for the guidance medium. Note that although $\overline{E}_w$ is in the direction of $\overline{J}_s$, $\sigma\overline{E}_w$ is not equal to $\overline{J}_s$ since $\overline{J}_s$ is a surface current, not a volume current. Dimensionally $\sigma\overline{E}_w$ has units amp/m^2 whereas $\overline{J}_s$ has units amp/m.

The dissipated power per unit length P_d is calculated by integrating the time-average power density $< \overline{E}_w \cdot \overline{J}_s^* >$ over the length along the guiding wall in a direction perpendicular to $\hat{z}$,

$$\begin{aligned} P_d &= \oint dl \; < \overline{E}_w \cdot \overline{J}_s^* > \\ &= \frac{1}{2} Re \left\{ \oint dl \sqrt{\frac{\omega\mu}{2\sigma}} (\hat{n} \times \overline{H}_w) \cdot (\hat{n} \times \overline{H}_w^*) \right\} \\ &= \frac{1}{2}\sqrt{\frac{\omega\mu}{2\sigma}} \oint dl \; |\overline{H}_w|^2 \end{aligned} \tag{33}$$

where $\overline{H}_w$ is the unperturbed tangential magnetic field component on the wall surface. We have used the identity

$$\begin{aligned} (\hat{n} \times \overline{H}_w) \cdot (\hat{n} \times \overline{H}_w^*) &= \hat{n} \cdot \left[\overline{H}_w \times (\hat{n} \times \overline{H}_w^*) \right] \\ &= \hat{n} \cdot \left[\hat{n} \, |\overline{H}_w|^2 - (\hat{n} \cdot \overline{H}_w) \overline{H}_w^* \right] \end{aligned}$$

and the fact that $\hat{n} \cdot \overline{H}_w^* = 0$ because $\overline{H}_w$ is perpendicular to the surface normal $\hat{n}$.

We can also derive the above result for P_d by the following arguments. Since a guided wave can be viewed as plane waves bouncing at the wall surfaces, a plane wave incident at the surface of a conducting medium results in power dissipation into the conductor. The transmitted plane wave is almost perpendicular to the surface regardless of the angles of incidence. The time-average Poynting power density flow into the conductor is $-\hat{n} \cdot < \overline{E}_w \times \overline{H}_w^* >$. Integrating over the length along the guiding wall in a direction perpendicular to $\hat{z}$, we obtain (33).

We now consider TM_m modes with the exact solutions

$$\begin{aligned}\overline{H} &= \hat{y} H_m \cos\frac{m\pi x}{d}\, e^{ik_z z} \\ \overline{E} &= H_m \left(\hat{x}\frac{k_z}{\omega\epsilon}\cos\frac{m\pi x}{d} - \hat{z}\frac{im\pi}{\omega\epsilon d}\sin\frac{m\pi x}{d}\right) e^{ik_z z}\end{aligned} \tag{34}$$

The time-average power flow in the waveguide P_f is

$$\begin{aligned}P_f &= \int_0^d dx \int_0^w dy\, \frac{1}{2}\frac{k_z}{\omega\epsilon}|H_m|^2 \cos^2\frac{m\pi x}{d} \\ &= \frac{wd}{4}\frac{k_z}{\omega\epsilon}(1+\delta_{0m})\,|H_m|^2\end{aligned} \tag{35}$$

where δ_{0m} is the Kronecker delta function with $\delta_{0m} = 0$ for $m \neq 0$ and $\delta_{00} = 1$. The time-average power dissipation per unit length P_d is

$$P_d = 2\int_0^w dy\, \frac{1}{2}|H_m|^2\sqrt{\frac{\omega\mu}{2\sigma}} = w\sqrt{\frac{\omega\mu}{2\sigma}}\,|H_m|^2 \tag{36}$$

where the factor 2 in front of the integral accounts for the fact that there are two walls at $x = 0$ and $x = d$ with equal magnitudes for H_w. The attenuation constant α is thus

$$\alpha^{\mathrm{TM}} = \frac{P_d}{2P_f} = \frac{1}{d}\sqrt{\frac{\omega\epsilon}{2\sigma}}\,\frac{2/(1+\delta_{0m})}{\sqrt{1-(k_{cm}/k)^2}} \tag{37}$$

Notice that for TEM mode, or, equivalently, the TM_0 mode, the corresponding attenuation constant is

$$\alpha^{\mathrm{TEM}} = \frac{1}{d}\sqrt{\frac{\omega\epsilon}{2\sigma}} \tag{38}$$

Following a similar procedure, we obtain the attenuation constants for TE modes

$$\alpha^{\mathrm{TE}} = \frac{1}{d}\sqrt{\frac{\omega\epsilon}{2\sigma}}\,\frac{2(k_{cm}/k)^2}{\sqrt{1-(k_{cm}/k)^2}} \tag{39}$$

In Figure 3.4.9 the attenuation constants for TEM, TM, and TE modes are plotted. The curve shapes for all TE_m and TM_m modes are the same if we let $\omega_c = m\pi/\sqrt{\mu\epsilon}\,d$. We note that for the TEM

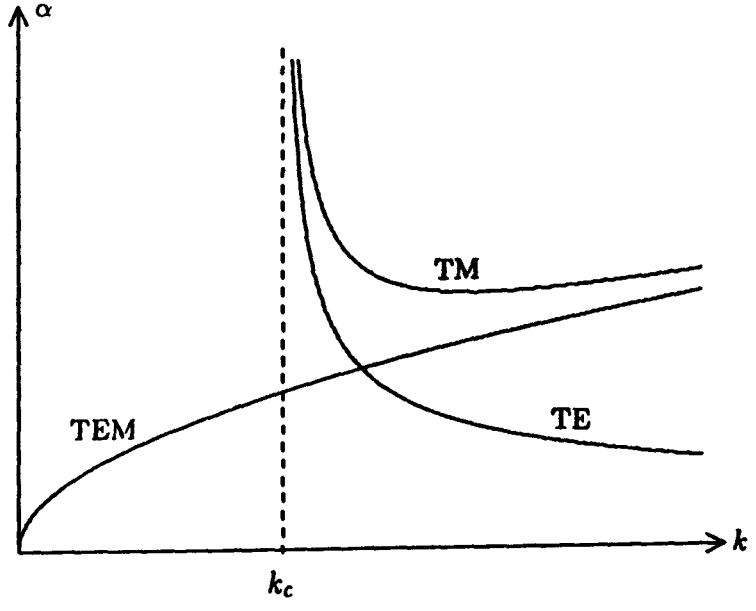


Figure 3.4.9 Attenuation constants.

mode α increases as the square root of ω. For TE modes α decreases monotonically with increasing frequency.

3.5 Guided Waves in Layered Media

Consider a layered medium as shown in Figure 3.5.1. We shall study wave guidance in region 0 by following the reflection coefficient formalism illustrated in 3.3. We write the TE wave solution in any region l

$$E_{ly} = \left(A_l e^{ik_{lz}x} + B_l e^{-ik_{lz}x}\right) e^{ik_z z} \tag{1}$$

The boundary condition of continuity of tangential components of $\overline{E}$ and $\overline{H}$ at $x = d_l$ yields

$$A_{l-1} e^{ik_{(l-1)z}d_l} + B_{l-1} e^{-ik_{(l-1)z}d_l} = A_l e^{ik_{lz}d_l} + B_l e^{-ik_{lz}d_l} \tag{2}$$

$$A_{l-1} e^{ik_{(l-1)z}d_l} - B_{l-1} e^{-ik_{(l-1)z}d_l} = p_{(l-1)l}\left[A_l e^{ik_{lz}d_l} - B_l e^{-ik_{lz}d_l}\right] \tag{3}$$

where

$$p_{(l-1)l} = \frac{\mu_{l-1} k_{lx}}{\mu_l k_{(l-1)x}} \tag{4}$$

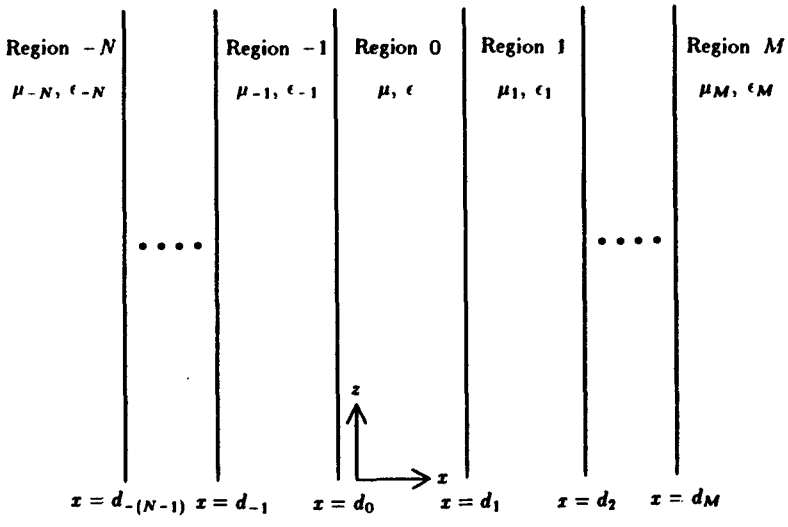


Figure 3.5.1 Guided waves in layered medium.

For TM waves we have

$$p_{(l-1)l} = \frac{\epsilon_{l-1} k_{lx}}{\epsilon_l k_{(l-1)x}} \tag{5}$$

Notice that $p_{l(l-1)} = 1/p_{(l-1)l}$. Solving A_{l-1} and B_{l-1} in terms of A_l and B_l, we find

$$A_{l-1} e^{ik_{(l-1)x} d_l} = \frac{1}{2}\left(1 + p_{(l-1)l}\right)\left[A_l e^{ik_{lx} d_l} + R_{(l-1)l} B_l e^{-ik_{lx} d_l}\right] \tag{6}$$

$$B_{l-1} e^{-ik_{(l-1)x} d_l} = \frac{1}{2}\left(1 + p_{(l-1)l}\right)\left[R_{(l-1)l} A_l e^{ik_{lx} d_l} + B_l e^{-ik_{lx} d_l}\right] \tag{7}$$

Solving A_l and B_l in terms of A_{l-1} and B_{l-1}, we find

$$A_l e^{ik_{lx} d_l} = \frac{1}{2}\left(1 + p_{l(l-1)}\right)\left[A_{l-1} e^{ik_{(l-1)x} d_l} + R_{l(l-1)} B_{l-1} e^{-ik_{(l-1)x} d_l}\right] \tag{8}$$

$$B_l e^{-ik_{lx} d_l} = \frac{1}{2}\left(1 + p_{l(l-1)}\right)\left[R_{l(l-1)} A_{l-1} e^{ik_{(l-1)x} d_l} + B_{l-1} e^{-ik_{(l-1)x} d_l}\right] \tag{9}$$

where

$$R_{(l-1)l} = \frac{1 - p_{(l-1)l}}{1 + p_{(l-1)l}} \tag{10}$$

and

$$R_{l(l-1)} = -R_{(l-1)l} \tag{11}$$

are the Fresnel reflection coefficients between regions $l-1$ and l. In region 0, we define reflection coefficients R_+ and R_- due to the stratified medium for $x > 0$ and $x < 0$ respectively as follows:

$$B_0 = R_+ A_0 \tag{12}$$

$$A_0 = R_- B_0 \tag{13}$$

The reflection coefficients R_+ and R_- can be determined from (6)–(9) in terms of continued fractions. From (6) and (7), we find

$$R_+ = \frac{B_0}{A_0} = \frac{e^{i2k_{1z}d_1}}{R_{01}} + \frac{[1-(1/R_{01}^2)]e^{i2(k_{1z}+k_z)d_1}}{(1/R_{01})e^{i2k_{1z}d_1} + (B_1/A_1)} \tag{14}$$

where B_1/A_1 can be expressed in terms of B_2/A_2 and so on until region M where $B_M/A_M = 0$. From (8) and (9) we find

$$\begin{aligned} R_- &= \frac{A_0}{B_0} \\ &= \frac{e^{-i2k_{z}d_0}}{R_{0(-1)}} + \frac{[1-(1/R_{0(-1)}^2)]e^{-i2(k_{-1z}+k_z)d_0}}{(1/R_{0(-1)})e^{-i2k_{-1z}d_0} + (A_{-1}/B_{-1})} \end{aligned} \tag{15}$$

where A_{-1}/B_{-1} can be expressed in terms of A_{-2}/B_{-2} and so on until region $-N$ where $A_{-N}/B_{-N} = 0$.

From (12) and (13), we obtain

$$R_+ R_- = 1 \tag{16}$$

which is the guidance condition for the guided waves in region 0.

We shall make use of the general formalism for guided waves in layered media given above to study several special cases. In this section we specialize to the case of wave guidance between two parallel plate conductors. We let $d_0 = 0$. The reflection coefficients become, for TE waves,

$$R_+^{\mathrm{TE}} = -e^{i2k_{z}d_1} \tag{17}$$

$$R_-^{\mathrm{TE}} = -1 \tag{18}$$

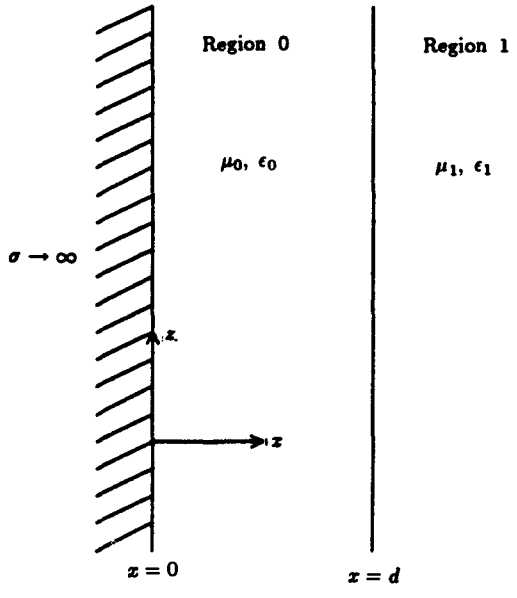


Figure 3.5.2 Guidance by an isotropic-medium-coated conductor.

The guidance condition (16) thus gives

$$k_x d_1 = m\pi \tag{19}$$

For TM waves, we have

$$R_+^{\mathrm{TM}} = e^{i2k_x d_1} \tag{20}$$

$$R_-^{\mathrm{TM}} = 1 \tag{21}$$

and the guidance condition is identical to (19).

The TE field solution for E_{0y} yields

$$\begin{aligned} E_{0y} &= B_0 \left[R_-^{\mathrm{TE}} e^{ik_x x} + e^{-ik_x x} \right] e^{ik_z z} \\ &= -i2B_0 \sin k_x x \, e^{ik_z z} \end{aligned}$$

The TM field solution for H_{0y} yields

$$\begin{aligned} H_{0y} &= A_0 \left[e^{ik_x x} + R_-^{\mathrm{TM}} e^{-ik_x x} \right] e^{ik_z z} \\ &= 2A_0 \cos k_x x \, e^{ik_z z} \end{aligned}$$

These solutions have been studied in detail in previous sections.

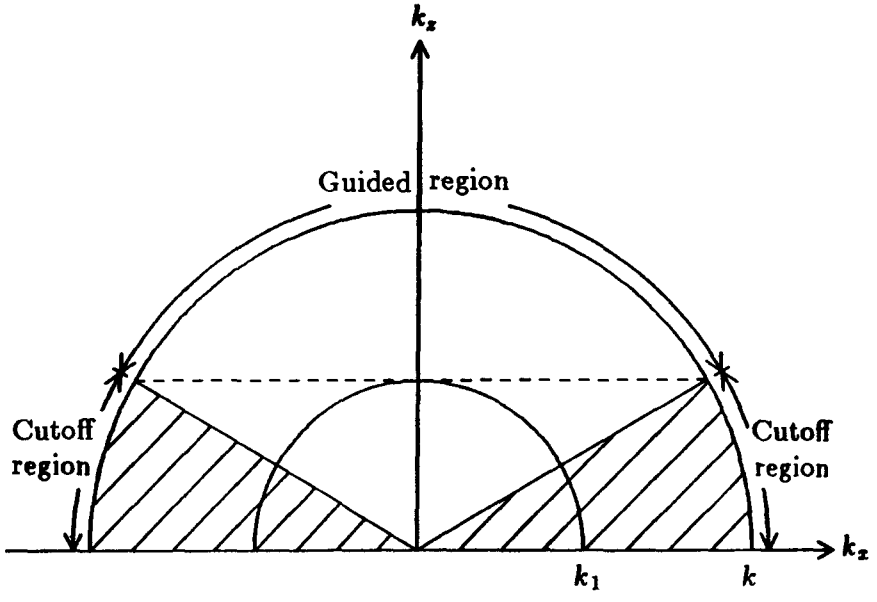


Figure 3.5.3 Guided and cutoff regions.

a. Guided Waves in Isotropic-Medium-Coated Conductor

Consider an isotropic-medium-coated conductor as shown in Figure 3.5.2. For the field to be guided in region 0, the field in region 1 must be evanescent in the $\hat{x}$ direction. We write the field solutions for the TE modes as follows:

$$\overline{E}_1 = \hat{y} E_1 e^{-\alpha_{1x} x} e^{ik_z z} \qquad d \leq x \qquad (22a)$$

$$\overline{E}_0 = \hat{y} \left[A_0 e^{ik_x x} + B_0 e^{-ik_x x} \right] e^{ik_z z} \qquad 0 \leq x \leq d \qquad (22b)$$

The dispersion relations for the two regions are

$$k_z^2 - \alpha_{1x}^2 = \omega^2 \mu_1 \epsilon_1 = k_1^2 \qquad (23a)$$

$$k_x^2 + k_z^2 = \omega^2 \mu \epsilon = k^2 \qquad (23b)$$

In Figure 3.5.3 we plot two k surfaces with radii k_1 and k for positive k_z as vertical axis and k_x as horizontal axis. For the wave to be guided, we must have k larger than k_1. When k_z becomes smaller than k_1, the wave in region 1 will no longer be evanescent because α_x becomes imaginary and the wave begins to propagate in the $\hat{x}$ direction. Thus the wave in region 0 is guided only when $k_z > k_1$.

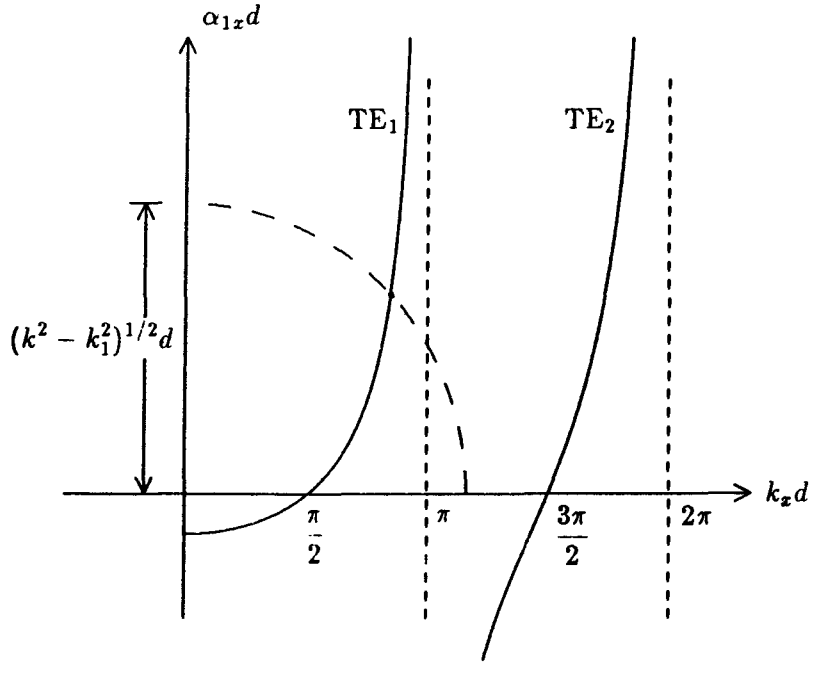


Figure 3.5.4 Interpretation of guidance condition for TE modes.

Let $B_0/A_0 = R_{0+}$ and $A_0/B_0 = R_{0-}$. The guidance condition states that

$$R_{0+}R_{0-} = 1 \tag{24}$$

The reflection coefficients R_{0+} and R_{0-} are found to be

$$R_{0+} = \frac{1 - p_{01}}{1 + p_{01}} e^{i2k_xd} \tag{25}$$

$$R_{0-} = -1 \tag{26}$$

where

$$p_{01} = \frac{i\mu\alpha_{1x}}{\mu_1 k_x}$$

for TE waves. From (26) we see that $A_0/B_0 = -1$, amounting to a phase shift of π when we view the incidence of a plane wave with amplitude B_0 and a reflected wave with amplitude A_0. From (25) we find $B_0/A_0 = R_{01}e^{i2k_xd}$, which can be interpreted as a plane wave reflected at the boundary $x = d$ with total reflection where $R_{01} =$

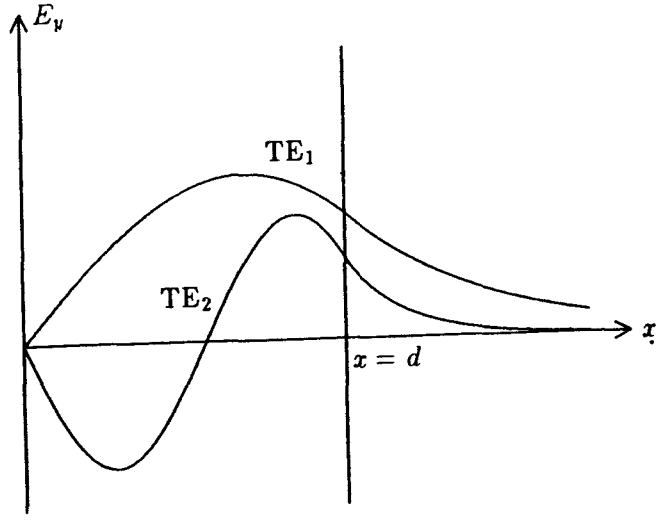


Figure 3.5.5 Electric field amplitudes for TE modes.

$e^{i2\phi_{01}}$, and $2\phi_{01} = -2\tan^{-1}(\mu\alpha_{1x}/\mu_1 k_x)$ is the Goos–Hänchen shift. We thus obtain the guidance condition $-R_{01}e^{i2k_x d} = 1$ which gives

$$2k_x d + 2\phi_{01} + \pi = 2m\pi \tag{27}$$

with $m = 1,2,\ldots$ It states that the total phase shift in the transverse $\hat{x}$ direction must be integer multiples of 2π to ensure constructive interference.

By using the expression for the Goos–Hänchen shift

$$2\phi_{01} = -2\tan^{-1}(\mu\alpha_{1x}/\mu_1 k_x)$$

we can express the guidance condition as

$$\alpha_{1x}d = -\frac{\mu_1}{\mu}k_x d\cot k_x d \tag{28}$$

We plot in Figure 3.5.4 the guidance condition on the two-dimensional plane with $k_x d$ as the horizontal axis and $\alpha_{1x}d$ as the vertical axis. The transverse wavenumber k_x for TE modes can be determined graphically from the dispersion relations (23a) and (23b) which give $k_x^2 + \alpha_{1x}^2 = k^2 - k_1^2$. This is a circle on the $\alpha_{1x}d - k_x d$ plane. We

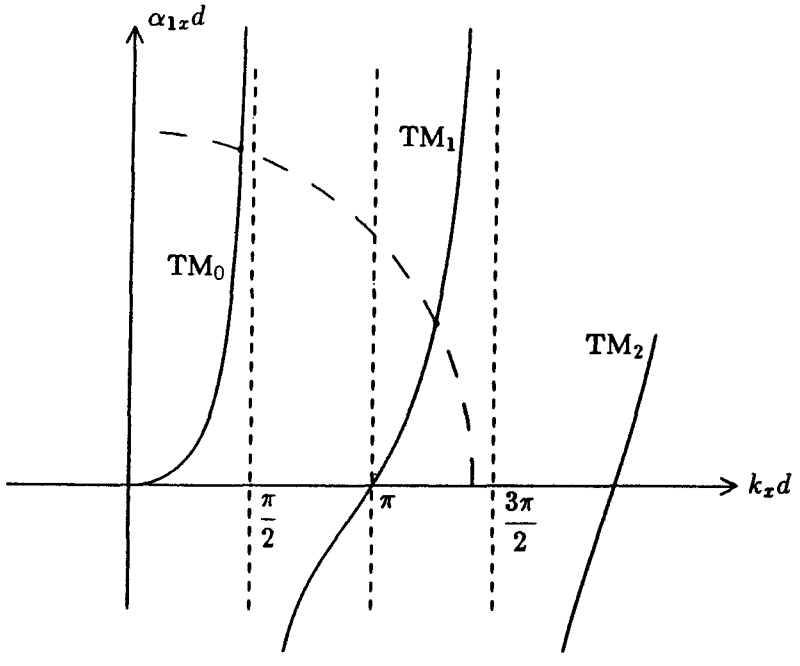


Figure 3.5.6 Interpretation of guidance condition for TM modes.

find that for TE_m mode, $(2m-1)\pi/2 < k_x d < m\pi$. The electric field in region 0 is found from (22) by using the guidance condition of $A_0/B_0 = -1$ and by matching the boundary condition at $x = d$,

$$\overline{E}_0 = \hat{y}\left(\frac{i}{2}E_1 e^{-\alpha_{1x}d}\right)\frac{\sin k_x x}{\sin k_x d}\, e^{ik_z z}$$

We show a field plot of E_y for TE_1 and TE_2 mode in Figure 3.5.5.

The waves are guided only when $k_z > k$ as shown in Figure 3.5.4. Cutoff occurs when α_{1x} becomes imaginary and the wave is no longer evanescent for $x > d$. The cutoff wavenumber is determined from $\phi_{01} = \alpha_{1x} = 0$ when $k_z = k_1$. Using the guidance condition (27) and (23*b*) we find $(2m-1)\pi/2 = k_x d = (k^2 - k_1^2)^{1/2} d$. We thus obtain

$$k_{cm} = \frac{(2m-1)\pi}{2d\sqrt{1-\mu_1\epsilon_1/\mu\epsilon}} \qquad m = 1, 2, \ldots \tag{29}$$

We see that the TE mode can have a large cutoff wavenumber k_c when $\mu_1\epsilon_1$ is very close to $\mu\epsilon$.

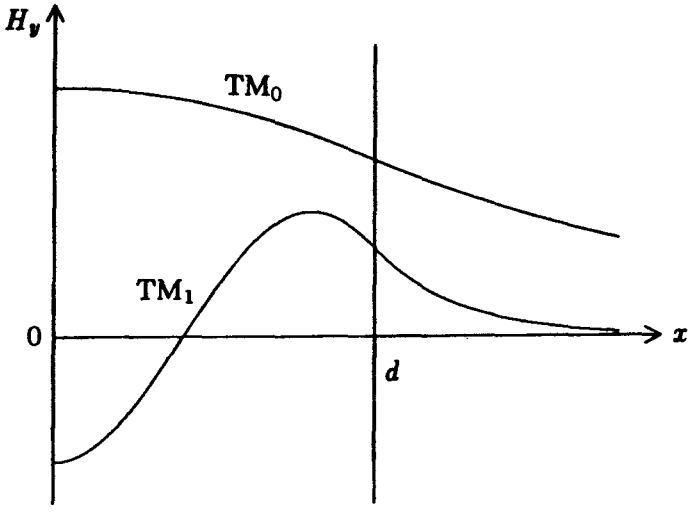


Figure 3.5.7 Magnetic field amplitudes for TM modes.

By the same token, the field solutions for the TM modes can be written as

$$\overline{H}_1 = \hat{y} H_1 e^{-\alpha_{1x} x} e^{ik_z z} \qquad d \leq x \qquad (30a)$$

$$\overline{H}_0 = \hat{y} \left[A_0 e^{ik_x x} + B_0 e^{-ik_x x} \right] e^{ik_z z} \qquad 0 \leq x \leq d \qquad (30b)$$

The guidance condition of $R_{0+}R_{0-} = 1$, with the fact that $R_{0-} = 1$, $R_{0+} = (1 - p_{01})/(1 + p_{01})$, and $p_{01} = i\epsilon\alpha_{1x}/\epsilon_1 k_x$, now gives

$$2k_x d + 2\phi_{01} = 2m\pi \qquad (31)$$

with $\phi_{01} = -\tan^{-1}(\epsilon\alpha_{1x}/\epsilon_1 k_x)$. Written in a form similar to (28), we have

$$\alpha_{1x} d = \frac{\epsilon_1}{\epsilon} k_x d \tan k_x d \qquad (32)$$

Figure 3.5.6 illustrates the graphical solution for k_x. We see that for the TM_m mode, $m\pi < k_x d < (2m + 1)\pi/2$. The magnetic field in region 0 is found from (30*b*) by using the guidance condition of $A_0/B_0 = 1$ and by matching the boundary condition at $x = d$. We find

$$H_y = \left(\frac{1}{2} H_0 e^{-ik_{1x} d} \right) \frac{\cos k_x x}{\cos k_x d} e^{ik_z z}$$

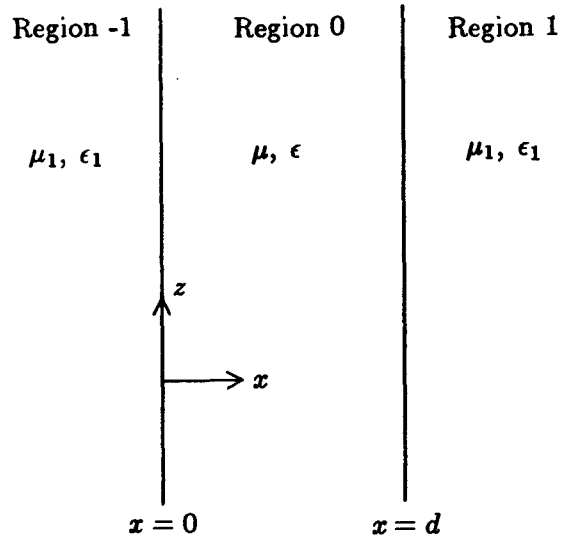


Figure 3.5.8 Guidance by a slab waveguide.

The field plots for H_y are illustrated in Figure 3.5.7. The cutoff wavenumber k_{cm} for TM_m modes is seen to be

$$k_{cm} = \frac{m\pi}{d\sqrt{1 - \mu_1\epsilon_1/\mu\epsilon}} \qquad m = 0, 1, 2, \ldots \tag{33}$$

As in the case of the parallel-plate waveguide, the TM_0 mode still has zero cutoff wavenumber or, equivalently, infinite cutoff wavelength.

b. Guided Waves in a Slab Dielectric Waveguide

Consider a symmetric slab waveguide with boundaries at $x = 0$ and d as shown in Figure 3.5.8. We first consider TE wave solutions. For waves to be guided inside the slab, the solutions for E_{-1y} and E_{1y} outside the slab must be evanescent in the $-\hat{x}$ and $\hat{x}$ directions. We write the electric field vectors in the three regions as

$$\overline{E}_1 = \hat{y}E_1 e^{-\alpha_{1z}x + ik_z z} \qquad d \le x \tag{34}$$

$$\overline{E}_0 = \hat{y}\left(A_0 e^{ik_z x} + B_0 e^{-ik_z x}\right) e^{ik_z z} \qquad 0 \le x \le d \tag{35}$$

$$\overline{E}_{-1} = \hat{y}E_{-1}\, e^{\alpha_{1z}x + ik_z z} \qquad x \le 0 \tag{36}$$

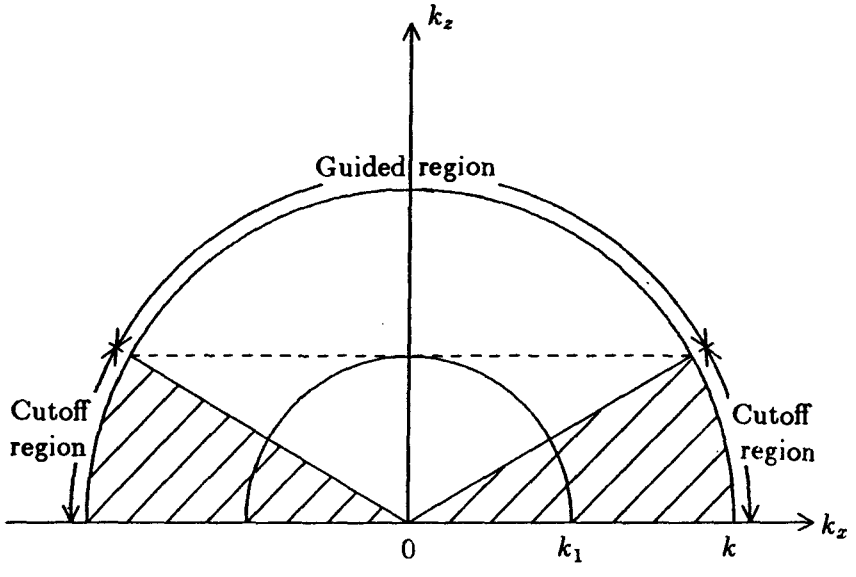


Figure 3.5.9 Guided and cutoff regions.

The solutions in regions −1 and 1 are both evanescent; $\overline{E}_1$ decays in $+\hat{x}$ direction and $\overline{E}_{-1}$ decays in $-\hat{x}$ direction. The dispersion relations in regions 0 and 1 are

$$k_z^2 - \alpha_{1x}^2 = \omega^2 \mu_1 \epsilon_1 = k_1^2 \tag{37}$$

$$k_z^2 + k_x^2 = \omega^2 \mu \epsilon = k^2 \tag{38}$$

In Figure 3.5.9 we plot two k surfaces with radii k_1 and k for positive k_z as vertical axis and k_x as horizontal axis. For the wave to be guided we must have k larger than k_1. When k_z becomes smaller than k_1, the wave in regions −1 and 1 will no longer be evanescent because α_{1x} becomes imaginary and the wave begins to propagate. The wave in region 0 is guided only when $k_z > k_1$.

Let $B_0/A_0 = R_{0+}$ and $A_0/B_0 = R_{0-}$. The guidance condition states that

$$R_{0+} R_{0-} = 1 \tag{39}$$

The reflection coefficients are

$$R_{0+} = \frac{1 - p_{01}}{1 + p_{01}} e^{i2k_z d} \tag{40}$$

$$R_{0-} = \frac{1 - p_{0(-1)}}{1 + p_{0(-1)}} \tag{41}$$

where $p_{01} = i\mu\alpha_{1x}/\mu_1 k_x = p_{0(-1)}$ for TE waves. From (40) and (41) we find $A_0/B_0 = (1-p_{01})/(1+p_{01}) = R_{01}$. Since $R_{01} = e^{i2\phi_{01}}$ and $2\phi_{01} = -2\tan^{-1}(\mu\alpha_{1x}/\mu_1 k_x)$ is the Goos–Hänchen shift at the boundary, we see that the ratio A_0/B_0 amounts to a phase shift of $2\phi_{01}$ at $x = 0$. At $x = d$, $B_0/A_0 = R_{01}e^{i2k_x d}$, which amounts to a phase shift of $2k_x d + 2\phi_{01}$. The guidance condition becomes $R_{01}^2 e^{i2k_x d} = e^{i2k_x d + i4\phi_{01}} = 1$, which yields

$$2k_x d + 4\phi_{01} = 2m\pi \tag{42}$$

Thus the total transverse phase shift in the $\hat{x}$ direction for a plane wave making a round trip adds to an integer multiple of 2π.

A graphical approach is useful in determining the propagation constant k_z for a symmetric slab waveguide. We rewrite the dispersion relations

$$(k_x d)^2 + (k_z d)^2 = (kd)^2 \tag{43}$$

$$(k_z d)^2 - (\alpha_{1x} d)^2 = (k_1 d)^2 \tag{44}$$

The guidance condition (42) can be written as

$$\alpha_{1x} d = \frac{\mu_1}{\mu} k_z d \tan\left(\frac{k_x d}{2} - \frac{m\pi}{2}\right) \tag{45}$$

For even m, we have $\alpha_{1x}d = (\mu_1 k_z d/\mu)\tan(k_x d/2)$ and for odd m, we have $\alpha_{1x}d = -(\mu_1 k_z d/\mu)\cot(k_x d/2)$. The guidance condition can be plotted on a two-dimensional plane determined by $\alpha_{1x}d$ and $k_x d$ [Fig. 3.5.10]. Eliminating k_z from the dispersion relations (43) and (44) we also obtain a family of circles on the $\alpha_{1x}d - k_z d$ plane.

$$(k_z d)^2 + (\alpha_{1x} d)^2 = (k^2 - k_1^2)d^2 \tag{46}$$

The two sets of curves intersect and give rise to values of α_{1x} and k_x, which in turn determine k_z. We observe that the larger the wavenumber k, the larger the radius, the more intersections, and hence, more modes can be guided. Notice that we must have $\alpha_{1x}d > 0$ in order for the wave to exponentially decay rather than grow. Cutoff occurs when

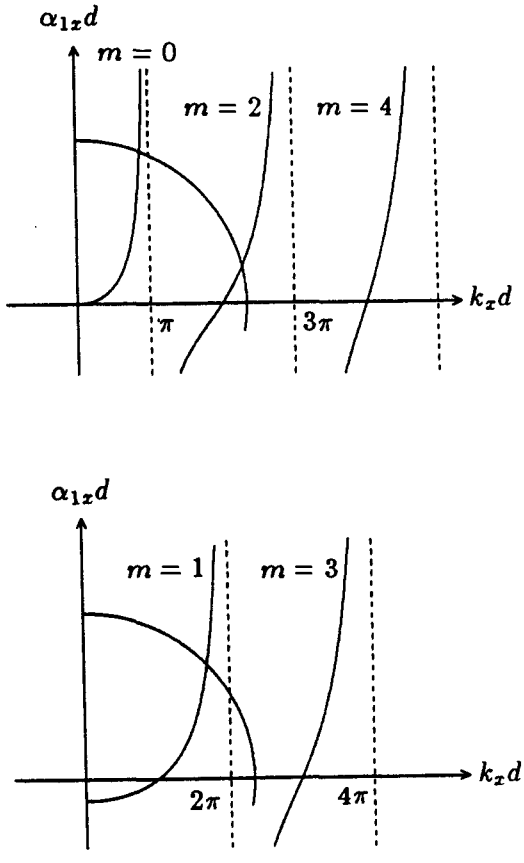


Figure 3.5.10 Interpretation of the guidance condition.

$\alpha_{1x} = 0$ and from (44) $k_z = k_1$. Making use of (43) and the guidance condition (42) with $\phi_{01} = 0$, we find the cutoff wavenumber

$$k_{cm} = \frac{m\pi}{d\sqrt{1 - \mu_1\epsilon_1/\mu\epsilon}} \tag{47}$$

Thus, the TE_0 and TM_0 modes possess zero cutoff wavenumber. For each value of k_x satisfying the guidance condition, the corresponding k_z is determined from (43).

We see that as k increases, more modes will be guided inside the slab waveguide. We wish to plot k_z as a function of k for the various modes. From the $k_z - k$ diagram, the phase and group velocities can be determined. The diagram can be generated graphically or numerically.

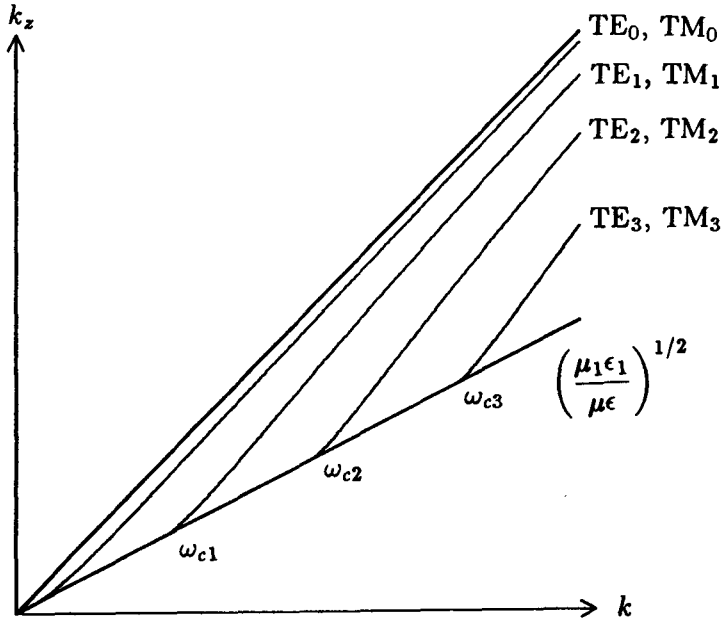


Figure 3.5.11 Propagation constant k_z as a function of k.

Let us examine in particular the asymptotic behavior of the $k_z - k$ diagram as $k \to \infty$ and as $k \to k_c$, where k_c is determined in (47).

As $k \to \infty$, we find from Figure 3.5.10 that $k_z = (m+1)\pi/d$, which is a constant. We can neglect $k_x d$ in (43) and find that

$$k_z = k \tag{48}$$

On the $k_z - k$ diagram this is a straight line with unit slope. As $k \to k_c$, $\alpha_x \to 0$, and we find from (44)

$$k_z = k\sqrt{\frac{\mu_1\epsilon_1}{\mu\epsilon}} \tag{49}$$

on the $k_z - k$ diagram this is a straight line with slope $(\mu_1\epsilon_1/\mu\epsilon)^{1/2}$. The group velocity of the guided waves as $k \to \infty$ is $(\mu\epsilon)^{-1/2}$, which is the velocity of light in the slab region. The group velocity as $k \to k_c$ is $(\mu_1\epsilon_1)^{-1/2}$, equal to the velocity of light in regions outside the slab. It can be shown by calculating $\partial k_z/\partial\omega$ that the group velocities for $k_c < k < \infty$ are smaller than $(\mu_1\epsilon_1)^{-1/2}$ and larger than $(\mu\epsilon)^{-1/2}$. The curves for the various modes are shown in Figure 3.5.11.

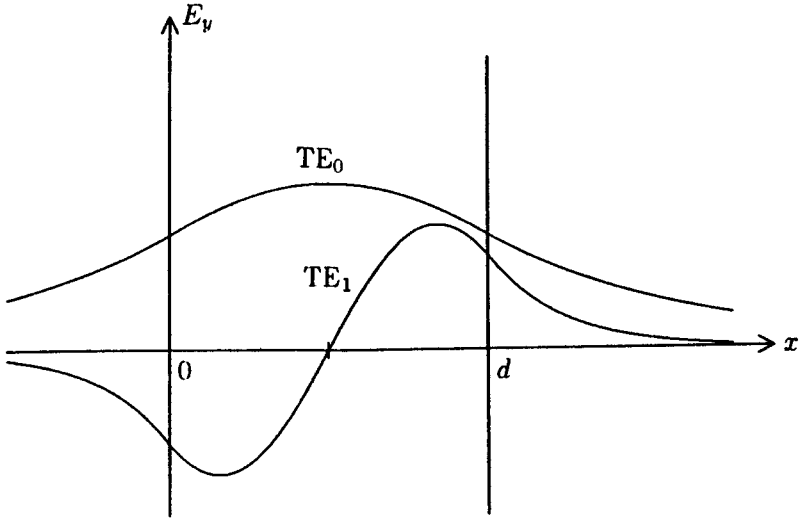


Figure 3.5.12 Field amplitudes for TE_0 and TE_1 modes.

The field solutions for the symmetric slab waveguide are, in view of (34)–(36) and the guidance condition,

$$\begin{aligned}\overline{E}_0 &= \hat{y}A_0\left[e^{ik_x x} + R_+^{\mathrm{TE}}e^{-ik_x x}\right] \\ &= \hat{y}A_0 e^{i\phi_{01}+ik_x d}\left[2\cos\left(k_x x - \frac{1}{2}k_x d - \frac{m\pi}{2}\right)\right]e^{ik_z z}\end{aligned} \tag{50}$$

In Figure 3.5.12, we plot the E_y field for the case of $m = 0$ and $m = 1$. Remember that $k_x d$ can be determined from Figure 3.5.10, where we see that $k_x d$ is larger than $m\pi$ and smaller than $(m+1)\pi$ for the mth mode. Thus, the higher the mode order, the more variations we shall have inside the slab region.

3.6 Cylindrical Waveguides

In treating guided waves along the $\hat{z}$ direction, the z dependence of all field vectors is written as $e^{\pm ik_z z}$ where k_z is the propagation constant and the $\pm$ signs indicate propagation along positive and negative $\hat{z}$ directions. With this dependence, we can replace $\partial^2/\partial z^2$ by $-k_z^2$. Due to its unique position in guided wave theory, the $\hat{z}$ direction is used to characterize guided modes. From the Maxwell equations, we

can express all field components parallel to the z axis. When all vectors are separated into their transverse and longitudinal components, Maxwell's two curl equations for isotropic media become

$$\left(\nabla_s + \hat{z}\frac{\partial}{\partial z}\right) \times \left(\overline{E}_s + \overline{E}_z\right) = i\omega\mu\left(\overline{H}_s + \overline{H}_z\right) \tag{1}$$

$$\left(\nabla_s + \hat{z}\frac{\partial}{\partial z}\right) \times \left(\overline{H}_s + \overline{H}_z\right) = -i\omega\epsilon\left(\overline{E}_s + \overline{E}_z\right) \tag{2}$$

where the subscript s denotes transverse components. Separating into transverse and longitudinal directions, we have

$$i\omega\mu\overline{H}_s = \nabla_s \times \overline{E}_z + \hat{z} \times \frac{\partial \overline{E}_s}{\partial z} \tag{3}$$

$$-i\omega\epsilon\overline{E}_s = \nabla_s \times \overline{H}_z + \hat{z} \times \frac{\partial \overline{H}_s}{\partial z} \tag{4}$$

$$i\omega\mu\overline{H}_z = \nabla_s \times \overline{E}_s \tag{5}$$

$$-i\omega\epsilon\overline{E}_z = \nabla_s \times \overline{H}_s \tag{6}$$

From (3) and (4) and making use of the identities $\hat{z}\times(\nabla_s\times\overline{E}_z) = \nabla_s E_z$ and $\hat{z} \times (\hat{z} \times \overline{E}_s) = -\overline{E}_s$, we can express $\overline{E}_s$ and $\overline{H}_s$ in terms of E_z and H_z

$$\overline{E}_s = \frac{1}{\omega^2\mu\epsilon - k_z^2}\left[\nabla_s\frac{\partial E_z}{\partial z} + i\omega\mu\nabla_s \times \overline{H}_z\right] \tag{7}$$

$$\overline{H}_s = \frac{1}{\omega^2\mu\epsilon - k_z^2}\left[\nabla_s\frac{\partial H_z}{\partial z} - i\omega\epsilon\nabla_s \times \overline{E}_z\right] \tag{8}$$

where we made use of the fact that $\partial^2/\partial z^2 = -k_z^2$ and $\overline{E}_z = \hat{z}E_z$ and $\overline{H}_z = \hat{z}H_z$. Substituting (7) in (5) and (8) in (6) we obtain

$$\left[\nabla_s^2 + \omega^2\mu\epsilon - k_z^2\right] E_z = 0 \tag{9}$$

$$\left[\nabla_s^2 + \omega^2\mu\epsilon - k_z^2\right] H_z = 0 \tag{10}$$

These are homogeneous Helmholtz equations for E_z and H_z. When the longitudinal components are solved from (9)–(10) and the transverse components determined from (7)–(8), we can proceed to match the appropriate boundary conditions imposed by the guiding structures.

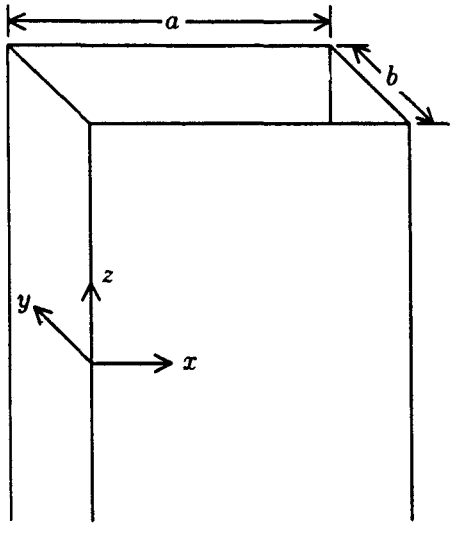


Figure 3.6.1 Metallic rectangular waveguides.

a. Rectangular Metallic Waveguides

Consider a metallic rectangular waveguide having dimensions a along the x axis and b along the y axis [Fig. 3.6.1]. We first investigate transverse magnetic (TM) fields where all magnetic fields are transverse to the direction of propagation $\hat{z}$. We have $H_z = 0$, and all field components can be derived from a single longitudinal component E_z. From (9) and considering the boundary conditions of vanishing tangential electric fields on the metallic wall surfaces, we obtain

$$E_z = \sin k_x x \sin k_y y \, e^{ik_z z} \tag{11}$$

The dispersion relation is

$$k_x^2 + k_y^2 + k_z^2 = \omega^2 \mu\epsilon = k^2 \tag{12}$$

The transverse components are found from (7)–(8),

$$E_x = \frac{ik_x k_z}{\omega^2 \mu\epsilon - k_z^2} \cos k_x x \sin k_y y \, e^{ik_z z} \tag{13}$$

$$E_y = \frac{ik_y k_z}{\omega^2 \mu\epsilon - k_z^2} \sin k_x x \cos k_y y \, e^{ik_z z} \tag{14}$$

$$H_x = \frac{-i\omega\epsilon k_y}{\omega^2\mu\epsilon - k_z^2} \sin k_x x \cos k_y y \, e^{ik_z z} \tag{15}$$

$$H_y = \frac{i\omega\epsilon k_x}{\omega^2\mu\epsilon - k_z^2} \cos k_x x \sin k_y y \, e^{ik_z z} \tag{16}$$

We see that at $x = 0$ and a, E_z and E_y vanish, and at $y = 0$ and b, E_z and E_x vanish, provided that

$$k_x a = m\pi \tag{17}$$

$$k_y b = n\pi \tag{18}$$

where m and n are integer numbers. Equations (17) and (18) are the guidance conditions. For TM waves neither m nor n can be zero because then E_z will be zero, too. Substituting the guidance conditions (17)–(18) in the field expressions, we see that for larger m there will be more variations for the fields as a function of x, and for larger n there will be more field variations along the $\hat{y}$ direction.

The dispersion relation (12) and the guidance conditions (17)–(18) combine to give the propagation constant

$$k_z = \sqrt{\omega^2\mu\epsilon - (m\pi/a)^2 - (n\pi/b)^2} \tag{19}$$

According to particular values of m and n, the TM waves inside the rectangular waveguide are classified into TM_{mn} modes. The first index m is associated with the number of variations along the $\hat{x}$ direction and the second index with the number of variations along the $\hat{y}$ direction.

Cutoff occurs when k_z becomes imaginary such that the wave attenuates exponentially along the direction of propagation. For a TM_{mn} mode, the cutoff wavenumber is

$$k_{cmn} = \sqrt{(m\pi/a)^2 + (n\pi/b)^2} \tag{20}$$

The lowest order TM mode is seen to be the TM_{11} mode. In Figure 3.6.2 we plot the propagation constant k_z as a function of k for the case $a = 2b$. As an example, if we let $a = 3$ cm, $b = 1.5$ cm, we find the cutoff wavenumber for the TM_{11} mode to be $k_{c11} = 234\ \mathrm{m}^{-1}$.

Next we examine TE fields which are derivable from a single longitudinal component H_z with $E_z = 0$. From (10) and considering

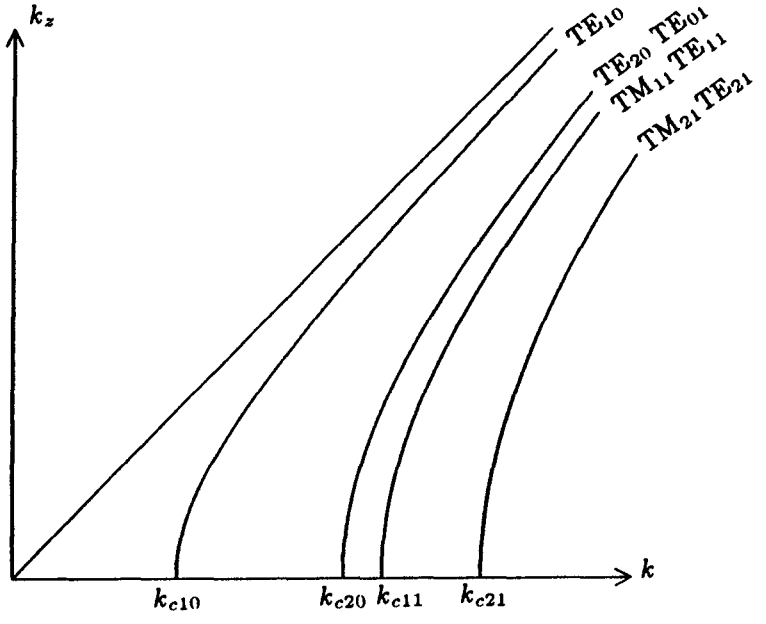


Figure 3.6.2 $k_z - k$ **diagram for guided modes.**

the boundary conditions of vanishing tangential electric fields on the metallic wall surfaces, we obtain

$$H_z = \cos k_x x \cos k_y y \, e^{ik_z z} \tag{21}$$

The dispersion relation is identical to (12). The transverse field components are found from (7)–(8):

$$H_x = \frac{-ik_x k_z}{\omega^2 \mu\epsilon - k_z^2} \sin k_x x \cos k_y y \, e^{ik_z z} \tag{22}$$

$$H_y = \frac{-ik_y k_z}{\omega^2 \mu\epsilon - k_z^2} \cos k_x x \sin k_y y \, e^{ik_z z} \tag{23}$$

$$E_x = \frac{-i\omega\mu k_y}{\omega^2 \mu\epsilon - k_z^2} \cos k_x x \sin k_y y \, e^{ik_z z} \tag{24}$$

$$E_y = \frac{i\omega\mu k_x}{\omega^2 \mu\epsilon - k_z^2} \sin k_x x \cos k_y y \, e^{ik_z z} \tag{25}$$

The guidance conditions are obtained from the boundary conditions of $E_x = 0$ at $y = 0$ and b and $E_y = 0$ at $x = 0$ and a. The result is

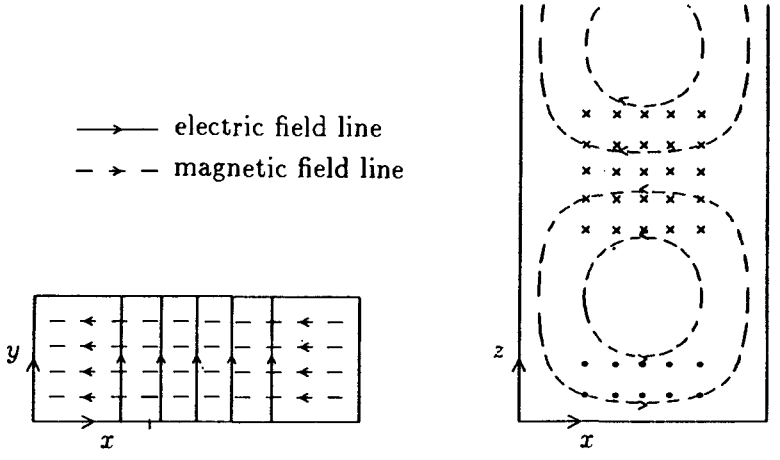


Figure 3.6.3 TE_{10} **mode in a rectangular waveguide.**

identical to (17) and (18). The propagation constant is again given as (19) and the cutoff wavenumbers are found to be (20).

There is, however, a very significant difference between the TE and TM modes. We noted below (18) that for TM_{mn} modes neither m nor n can be zero. For the TE_{mn} modes it is possible to have either m or n or both m and n equal to zero. For $m = n = 0$, we find $H_z = e^{ikz}$. The equation $\nabla \cdot \overline{H} = 0$ implies $k = \omega(\mu\epsilon)^{1/2} = 0$ and consequently the TM_{00} mode is a static field solution in the waveguide. We now argue that the TEM mode for which $E_z = H_z = 0$ cannot exist inside the hollow waveguide. From $\nabla \cdot \overline{H} = \nabla \cdot \overline{H}_s = 0$ and the boundary condition of vanishing normal $\overline{H}$ field on the wall, we find that the $\overline{H}$ field must form closed loops and $\nabla \times \overline{H}_s = \hat{z}J_z - \hat{z}i\omega\epsilon E_z \neq 0$. Thus either $J_z \neq 0$, which implies the waveguide is not hollow, or $E_z \neq 0$, which implies that the mode cannot be TEM. As a corollary, we have in fact shown that TEM modes do exist when there is another conductor to support the conduction current J_z, as in the case of a coaxial line.

Assuming $a > b$, we see from (20) that the lowest order TE mode will be TE_{10}, with cutoff wavenumber

$$k_{c10} = \pi/a \tag{26}$$

The field components of the TE_{10} mode are

$$H_z = \cos\frac{\pi x}{a} e^{ik_z z} \tag{27}$$

$$H_x = \frac{-ik_z a}{\pi} \sin \frac{\pi x}{a} e^{ik_z z} \tag{28}$$

$$E_y = \frac{i\omega\mu a}{\pi} \sin \frac{\pi x}{a} e^{ik_z z} \tag{29}$$

The electric field has only a y component. A field plot for the TE_{10} mode is shown in Figure 3.6.3. Written in the form of the superposition of two bouncing plane waves, (29) becomes

$$E_y = \frac{\omega\mu a}{2\pi} \left[e^{i\frac{\pi x}{a} + ik_z z} - e^{-i\frac{\pi x}{a} + ik_z z} \right] \tag{30}$$

Higher-order TE and TM modes can also be interpreted as plane waves bouncing around the four walls and propagating along $\hat{z}$ with the propagation constant k_z. The propagation constant k_z for the various modes are plotted in Figure 3.6.2 for the case of $a = 2b$. Since the TE_{10} mode has the lowest cutoff frequency, it is the fundamental mode or the dominant mode of the rectangular waveguide.

b. Circular Metallic Waveguides

In order to study cylindrical waveguides of circular cross-sections, we first consider the wave equation for E_z and H_z in cylindrical coordinates

$$\left[\frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2}{\partial \phi^2} + k_\rho^2 \right] \begin{Bmatrix} E_z \\ H_z \end{Bmatrix} = 0 \tag{31}$$

where $k_\rho^2 = \omega^2\mu\epsilon - k_z^2$. Solutions to the wave equation are Bessel functions multiplied by sinusoidal functions. The sinusoids can be combinations of $\sin m\phi$, $\cos m\phi$, or $e^{\pm im\phi}$. Substituting in (31) and making use of the transformation $\xi = k_\rho \rho$, we have the Bessel equation

$$\left[\frac{1}{\xi} \frac{d}{d\xi} \left(\xi \frac{d}{d\xi} \right) + \left(1 - \frac{m^2}{\xi^2} \right) \right] B(\xi) = 0 \tag{32}$$

This has solutions in the form of the Bessel function $J_m(\xi)$, Neumann function $N_m(\xi)$, Hankel function of the first kind $H_m^{(1)}(\xi)$, or Hankel function of the second kind $H_m^{(2)}(\xi)$. The two kinds of Hankel functions are related to the Bessel and Neumann functions in the following manner:

$$H_m^{(1)}(\xi) = J_m(\xi) + iN_m(\xi) \tag{33}$$

$$H_m^{(2)}(\xi) = J_m(\xi) - iN_m(\xi) \tag{34}$$

ξ / $B(\xi)$	$\xi \to 0$		$\xi \to \infty$
	$m = 0$	$Re\{m\} > 0$	
$J_m(\xi)$	1	$\dfrac{(\xi/2)^m}{\Gamma(m+1)}$	$\sqrt{2/\pi\xi}\cos(\xi - \dfrac{m\pi}{2} - \dfrac{\pi}{4})$
$N_m(\xi)$	$\dfrac{2}{\pi}\ln\xi$	$-\dfrac{\Gamma(m)}{\pi}(\dfrac{2}{\xi})^m$	$\sqrt{2/\pi\xi}\sin(\xi - \dfrac{m\pi}{2} - \dfrac{\pi}{4})$
$H_m^{(1)}(\xi)$	$i\dfrac{2}{\pi}\ln\xi$	$-i\dfrac{\Gamma(m)}{\pi}(\dfrac{2}{\xi})^m$	$\sqrt{2/\pi\xi}\exp[i(\xi - \dfrac{m\pi}{2} - \dfrac{\pi}{4})]$
$H_m^{(2)}(\xi)$	$-i\dfrac{2}{\pi}\ln\xi$	$i\dfrac{\Gamma(m)}{\pi}(\dfrac{2}{\xi})^m$	$\sqrt{2/\pi\xi}\exp[-i(\xi - \dfrac{m\pi}{2} - \dfrac{\pi}{4})]$

Table 3.6.1 Limiting values of J_m, N_m, $H_m^{(1)}$ and $H_m^{(2)}$.

The asymptotic values of these functions are listed in Table 3.6.1. As $\xi \to \infty$, J_m behaves as cosine, N_m as sine, and H_m as exponents. As $\xi \to 0$, all but J_m becomes singular.

Let $B_m(\xi)$ represent $J_m(\xi)$, $N_m(\xi)$, $H_m^{(1)}(\xi)$, or $H_m^{(2)}(\xi)$. The recurrence formulas for the Bessel functions $B_m(\xi)$ are as follows:

$$\begin{aligned} B'_m(\xi) &= B_{m-1}(\xi) - \frac{m}{\xi}B_m(\xi) \\ &= -B_{m+1}(\xi) + \frac{m}{\xi}B_m(\xi) \end{aligned} \tag{35}$$

where the prime on $B(\xi)$ indicates derivative with respect to its argument. In Figures 3.6.4–3.6.6 we plot $J_m(\xi)$, $J'_m(\xi)$, and $N_m(\xi)$ for $m = 0, 1,$ and 2.

Consider a circular metallic waveguide with radius a [Fig. 3.6.7]. The boundary conditions require that E_z and E_ϕ vanish at $\rho = a$. Since fields must not be singular at $\rho = 0$, the Bessel function J_m is the only logical solution. For TM waves, we have

$$E_z = J_m(k_\rho \rho)\begin{Bmatrix} \sin m\phi \\ \cos m\phi \end{Bmatrix} e^{ik_z z} \tag{36}$$

with the dispersion relation

$$k_z^2 + k_\rho^2 = \omega^2 \mu\epsilon \tag{37}$$

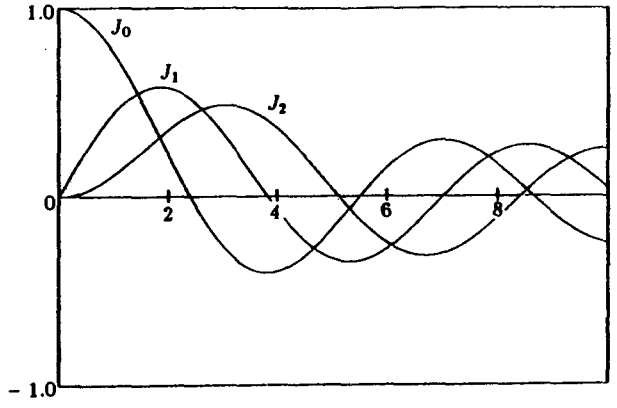


Figure 3.6.4 Bessel functions.

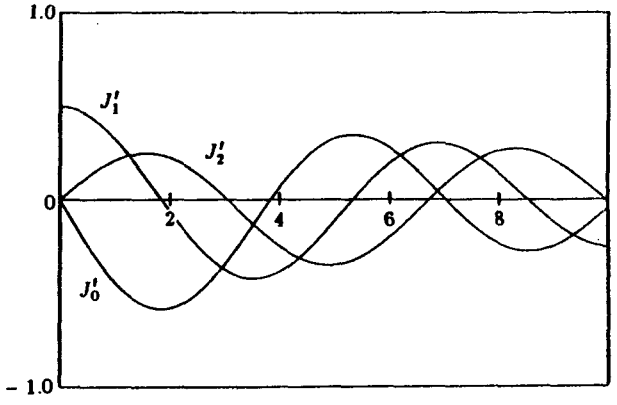


Figure 3.6.5 Derivatives of Bessel functions.

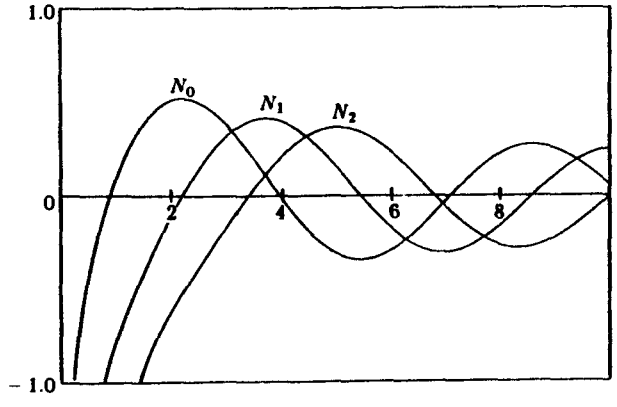


Figure 3.6.6 Neumann functions.

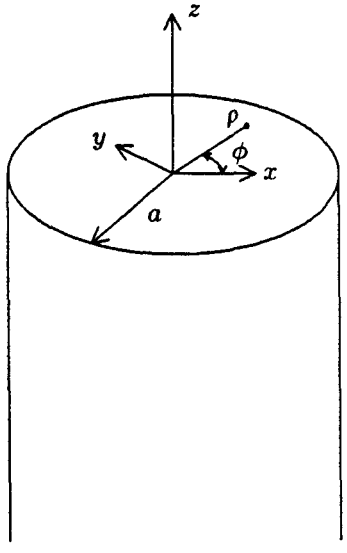


Figure 3.6.7 Circular metallic waveguides.

Transverse field components are obtained from the guided wave formalism (7)–(8),

$$E_\rho = \frac{ik_z k_\rho}{\omega^2\mu\epsilon - k_z^2} J_m'(k_\rho\rho)\begin{Bmatrix}\sin m\phi\\ \cos m\phi\end{Bmatrix} e^{ik_z z} \tag{38}$$

$$E_\phi = \frac{ik_z}{\omega^2\mu\epsilon - k_z^2}\frac{m}{\rho} J_m(k_\rho\rho)\begin{Bmatrix}\cos m\phi\\ -\sin m\phi\end{Bmatrix} e^{ik_z z} \tag{39}$$

$$H_\rho = \frac{-i\omega\epsilon}{\omega^2\mu\epsilon - k_z^2}\frac{m}{\rho} J_m(k_\rho\rho)\begin{Bmatrix}\cos m\phi\\ -\sin m\phi\end{Bmatrix} e^{ik_z z} \tag{40}$$

$$H_\phi = \frac{i\omega\epsilon k_\rho}{\omega^2\mu\epsilon - k_z^2} J_m'(k_\rho\rho)\begin{Bmatrix}\sin m\phi\\ \cos m\phi\end{Bmatrix} e^{ik_z z} \tag{41}$$

where the prime on the Bessel function denotes derivative with respect to its argument. The boundary condition of vanishing E_z and E_ϕ at $\rho = a$ gives the guidance condition

$$J_m(k_\rho a) = 0 \tag{42}$$

Let ξ_{mn} denote the n th root of the m th order Bessel function such that $J_m(\xi_{mn}) = 0$, we find from the guidance condition (42) and the

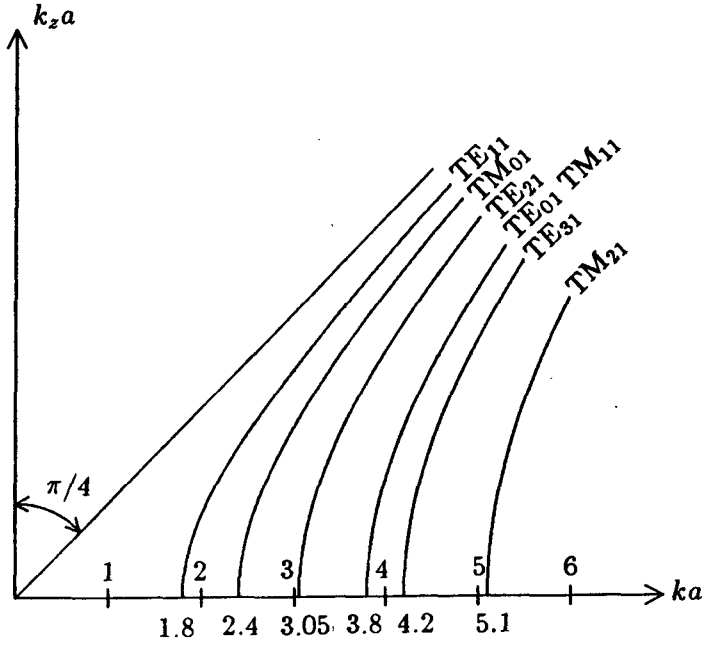


Figure 3.6.8 $k_z a - ka$ **diagrams for guided modes.**

dispersion relation (37)

$$k_z = \sqrt{\omega^2 \mu\epsilon - (\xi_{mn}/a)^2} \tag{43}$$

For each value of ξ_{mn} we label the wave solution TM_{mn} mode, where m is associated with the number of variations in the $\hat{\phi}$ direction and n with the number of variations in the $\hat{\rho}$ direction. The cutoff wavenumber for the TM_{mn} modes are

$$k_{cmn} = \xi_{mn}/a \tag{44}$$

We find from Figure 3.6.4 that for the TM_{01} mode $\xi_{01} = 2.4$ and for the TM_{11} mode $\xi_{11} = 3.8$. The $k_z a - ka$ diagram for the guided TM_{mn} modes is plotted in Figure 3.6.8.

For TE wave solutions we have

$$H_z = J_m(k_\rho \rho) \begin{Bmatrix} \sin m\phi \\ \cos m\phi \end{Bmatrix} e^{ik_z z} \tag{45}$$

with dispersion relation

$$k_z^2 + k_\rho^2 = \omega^2 \mu\epsilon \tag{46}$$

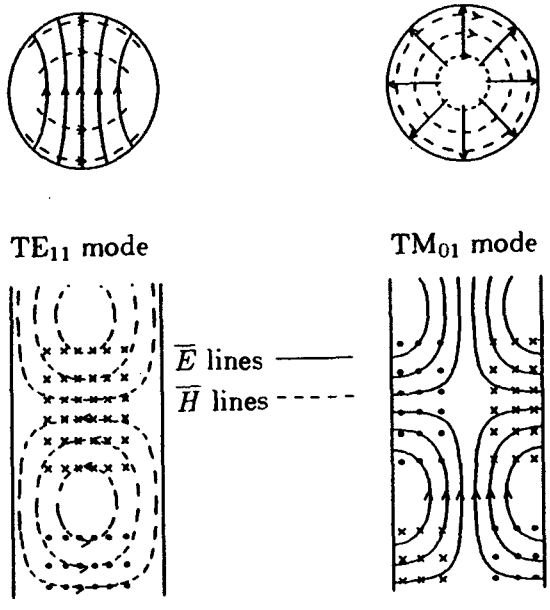


Figure 3.6.9 Field lines for TE_{11} and TM_{01} modes.

The transverse field components are

$$H_\rho = \frac{ik_z k_\rho}{\omega^2\mu\epsilon - k_z^2} J_m'(k_\rho\rho)\begin{Bmatrix}\sin m\phi\\ \cos m\phi\end{Bmatrix} e^{ik_z z} \tag{47}$$

$$H_\phi = \frac{ik_z}{\omega^2\mu\epsilon - k_z^2}\frac{m}{\rho} J_m(k_\rho\rho)\begin{Bmatrix}\cos m\phi\\ -\sin m\phi\end{Bmatrix} e^{ik_z z} \tag{48}$$

$$E_\rho = \frac{i\omega\mu}{\omega^2\mu\epsilon - k_z^2}\frac{m}{\rho} J_m(k_\rho\rho)\begin{Bmatrix}\cos m\phi\\ -\sin m\phi\end{Bmatrix} e^{ik_z z} \tag{49}$$

$$E_\phi = \frac{-i\omega\mu k_\rho}{\omega^2\mu\epsilon - k_z^2} J_m'(k_\rho\rho)\begin{Bmatrix}\sin m\phi\\ \cos m\phi\end{Bmatrix} e^{ik_z z} \tag{50}$$

The boundary condition of vanishing E_z and E_ϕ at $\rho = a$ gives the guidance condition

$$J_m'(k_\rho a) = 0 \tag{51}$$

Letting ξ_{mn}' denote the roots of J_m' such that $J_m'(\xi_{mn}') = 0$, we find from the dispersion relation (46) and the guidance condition (51)

$$k_z = \sqrt{\omega^2\mu\epsilon - (\xi_{mn}'/a)^2} \tag{52}$$

For each value of ξ'_{mn} we label the wave solution TE_{mn} mode where m is associated with the number of variations in the $\hat{\phi}$ direction and n with the number of variations in the $\hat{\rho}$ direction. The field patterns for the TE_{11} and TM_{01} modes are shown in Figure 3.6.9. The cutoff frequencies for the TE_{mn} modes are

$$k_{cmn} = \xi'_{mn}/a \tag{53}$$

We see from Figure 3.6.5 that for the TE_{11} mode $\xi'_{11} = 1.8$, for the TE_{21} mode $\xi'_{21} = 3.05$, and for the TE_{01} mode $\xi'_{01} = 3.8$. The $k_z a - ka$ diagram for the guided TE_{mn} modes is shown in Figure 3.6.8.

c. *Circular Dielectric Waveguides*

Consider a circular waveguide of radius a made of an isotropic medium with constitutive parameters μ and ϵ embedded in another isotropic medium characterized by μ_1 and ϵ_1 [Fig. 3.6.10]. For dielectric waveguides, $\mu_1 = \mu$. But for general discussions we shall keep $\mu_1 \neq \mu$ unless otherwise specified. As we shall see, in order to satisfy the boundary conditions of continuity of tangential $\overline{E}$ and $\overline{H}$ fields at $\rho = a$, the solutions in general require a combination of both TE and TM waves. The guided modes with both E_z and H_z components present are known as hybrid modes.

Inside the waveguide, solutions for E_z and H_z are Bessel functions,

$$E_z = AJ_m(k_\rho \rho)\cos m\phi\, e^{ik_z z} \tag{54}$$

$$H_z = BJ_m(k_\rho \rho)\sin m\phi\, e^{ik_z z} \tag{55}$$

with the dispersion relation

$$k_z^2 + k_\rho^2 = \omega^2 \mu\epsilon = k^2 \tag{56}$$

In (54) and (55), $\cos m\phi$ and $\sin m\phi$ are chosen for E_z and H_z in anticipation of matching the boundary conditions. They can also be replaced with $\sin m\phi$ and $\cos m\phi$, respectively.

Outside the waveguide, the field associated with the guided waves must be evanescent in the $\hat{\rho}$ direction. The proper choice of the solution will be Hankel functions with imaginary arguments, known as the modified Hankel functions. Letting $k_{1\rho} = i\alpha_{1\rho}$, we write for $\rho \geq a$

$$E_z = CH_m^{(1)}(i\alpha_{1\rho}\rho)\cos m\phi\, e^{ik_z z} \tag{57}$$

$$H_z = DH_m^{(1)}(i\alpha_{1\rho}\rho)\sin m\phi\, e^{ik_z z} \tag{58}$$

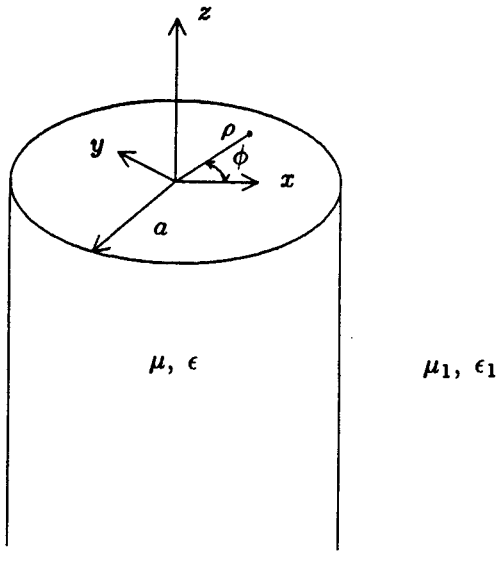


Figure 3.6.10 Circular dielectric waveguides.

with the dispersion relation

$$k_z^2 - \alpha_{1\rho}^2 = \omega^2 \mu_1 \epsilon_1 = k^2 \tag{59}$$

The choice of $\cos m\phi$ and $\sin m\phi$ for E_z and H_z in (57)–(58) facilitates their equality with (54)–(55) at $\rho = a$ for all azimuthal angles ϕ.

Transverse components for the fields are determined from (7)–(8). Matching the z and ϕ components at $\rho = a$ we obtain the following four equations for A, B, C, and D,

$$AJ_m(k_\rho a) = CH_m^{(1)}(i\alpha_{1\rho}a) \tag{60}$$

$$BJ_m(k_\rho a) = DH_m^{(1)}(i\alpha_{1\rho}a) \tag{61}$$

$$A\frac{\omega\epsilon}{k_\rho a}J_m'(k_\rho a) + B\frac{mk_z}{k_\rho^2 a^2}J_m(k_\rho a) = -iC\frac{\omega\epsilon_1}{\alpha_{1\rho}a}H_m^{(1)\prime}(i\alpha_{1\rho}a) - D\frac{mk_z}{\alpha_{1\rho}^2 a^2}H_m^{(1)}(i\alpha_{1\rho}a) \tag{62}$$

$$A\frac{mk_z}{k_\rho^2 a^2}J_m(k_\rho a) + B\frac{\omega\mu}{k_\rho a}J_m'(k_\rho a)$$

$$= -C\frac{mk_z}{\alpha_{1\rho}^2 a^2}H_m^{(1)}(i\alpha_{1\rho}a) - iD\frac{\omega\mu_1}{\alpha_{1\rho}a}H_m^{(1)\prime}(i\alpha_{1\rho}a) \quad (63)$$

Equations (60)–(61) result from the continuity of E_z and H_z; (62)–(63) result from the continuity of E_ϕ and H_ϕ. These four equations give rise to guidance conditions for the circular dielectric waveguide.

We observe that for TE waves $A = C = 0$, (60)–(63) reduce to two equations for B and D only when $m = 0$. Similarly, for TM waves to exist we must also have $m = 0$. In this case we obtain

$$\frac{J_0(k_\rho a)}{J_0'(k_\rho a)} = p_{10}\frac{H_0^{(1)}(i\alpha_{1\rho}a)}{H_0^{(1)\prime}(i\alpha_{1\rho}a)} \quad (64)$$

where

$$p_{10} = i\frac{\mu\alpha_{1\rho}}{\mu_1 k_\rho} \quad (65a)$$

for TE_{0p} modes and

$$p_{10} = i\frac{\epsilon\alpha_{1\rho}}{\epsilon_1 k_\rho} \quad (65b)$$

for TM_{0p} modes, where the subscript $0p$ denotes mode number arranged in order of increasing cutoff wavenumber.

Cutoff occurs when $\alpha_{1\rho}$ becomes imaginary such that the argument of the first-kind Hankel function becomes real and the field outside is radiating in the radial direction ρ. This cutoff criterion is identical to that used in the case of slab dielectric waveguides. Near cutoff, $\alpha_{1\rho} \to 0$, $k_z \to k_1$, and $k_\rho \to k_c(1 - \mu_1\epsilon_1/\mu\epsilon)$. We now determine the cutoff wavenumbers for the various guided modes.

In view of the asymptotic values for the Hankel functions as $\alpha_{1\rho} \to 0$, we see that (64) and (65) yield identical equations for TE and TM modes at cutoff,

$$J_0\left(k_{c0p}a\sqrt{1 - \frac{n_1^2}{n^2}}\right) = 0 \quad (66)$$

where $n_1 = c(\mu_1\epsilon_1)^{1/2}$ and $n = c(\mu\epsilon)^{1/2}$. We thus have for TE_{0p} and TM_{0p} modes $k_{c0p}a(1 - n_1^2/n^2)^{1/2} = \xi_{0p}$, where ξ_{0p} is the pth mode of the zeroth order Bessel function $J_0(\xi)$.

When $m \neq 0$, separation into TE and TM modes is no longer possible; the $m \neq 0$ modes are hybrid. Eliminating A, B, C, and

D from (60)–(63) we find

$$J_m^2(k_\rho a)H_m^{(1)^2}(i\alpha_{1\rho}a)\left\{m^2k_z^2\left(\frac{1}{k_\rho^2a^2}+\frac{1}{\alpha_{1\rho}^2a^2}\right)^2 + \frac{\omega^2\mu_1\epsilon_1}{\alpha_{1\rho}^2a^2}\left[p_{10}^{\mathrm{TE}}\frac{J_m'(k_\rho a)}{J_m(k_\rho a)}-\frac{H_m^{(1)'}(i\alpha_{1\rho}a)}{H_m^{(1)}(i\alpha_{1\rho}a)}\right] \cdot\left[p_{10}^{\mathrm{TM}}\frac{J_m'(k_\rho a)}{J_m(k_\rho a)}-\frac{H_m^{(1)'}(i\alpha_{1\rho}a)}{H_m^{(1)}(i\alpha_{1\rho}a)}\right]\right\}=0 \quad (67)$$

The hybrid modes are called EH modes when $J_m(k_\rho a)$ is set to zero. The hybrid modes are called HE modes when the guidance conditions are obtained by setting the term in braces of (67) equal to zero.

For the hybrid EH_{mp} modes, we see that as $\alpha_{1\rho}\to 0, k_\rho \to k_c(1-n_1^2/n^2)^{1/2}$, and $J_m(k_\rho a)=0$ leads to

$$k_{cmp}\left(1-\frac{n_1^2}{n^2}\right)^{1/2}=\xi_{mp} \qquad \text{with } \xi_{mp}\neq 0 \quad (68)$$

where ξ_{mp} is the pth root of the mth order Bessel function $J_m(\xi)$. The root of $\xi_{mp}=0$ is excluded and the EH modes do not have modes with zero cutoff wavenumber. This is because as $k_c\to 0$, both $\alpha_{1\rho}$ and k_ρ also approach zero and (67) cannot be satisfied.

For the hybrid HE_{mp} modes, we use the fact that $k_z^2=k^2-k_\rho^2=k_1^2+\alpha_{1\rho}^2$ and the recurrence relations for the Bessel and Hankel functions to obtain

$$m^2\left(\frac{k^2}{k_\rho^2a^2}+\frac{k_1^2}{\alpha_{1\rho}^2a^2}\right)\left(\frac{1}{k_\rho^2a^2}+\frac{1}{\alpha_{1\rho}^2a^2}\right) + \frac{\omega^2\mu_1\epsilon_1}{\alpha_{1\rho}^2a^2}\left(p_{10}^{\mathrm{TE}}\frac{J_m'(k_\rho a)}{J_m(k_\rho a)}-\frac{H_{m-1}^{(1)}(i\alpha_{1\rho}a)}{H_m^{(1)}(i\alpha_{1\rho}a)}-i\frac{m}{\alpha_{1\rho}a}\right) \cdot\left(p_{10}^{\mathrm{TM}}\frac{J_m'(k_\rho a)}{J_m(k_\rho a)}-\frac{H_{m-1}^{(1)}(i\alpha_{1\rho}a)}{H_m^{(1)}(i\alpha_{1\rho}a)}-i\frac{m}{\alpha_{1\rho}a}\right)=0 \quad (69)$$

When $m>1$, we use the asymptotic formula

$$\frac{H_{m-1}^{(1)}(i\alpha_{1\rho}a)}{H_m^{(1)}(i\alpha_{1\rho}a)}\sim\frac{i\alpha_{1\rho}a}{2(m-1)} \quad (70)$$

as $\alpha_{1\rho}a \to 0$. Keeping terms of the order of $1/\alpha_{1\rho}^2 a^2$ as the terms of the order of $1/\alpha_{1\rho}^4 a^4$ cancel each other, we obtain from (69)

$$\frac{\mu\epsilon_1 + \mu_1\epsilon}{k_\rho a}\frac{J_m'(k_\rho a)}{J_m(k_\rho a)} + m\frac{\mu_1\epsilon_1 + \mu\epsilon}{k_\rho^2 a^2} - \frac{\mu_1\epsilon_1}{m-1} = 0 \tag{71}$$

This equation determines cutoff frequencies for HE_{mp} modes with $m > 1$. For $m = 1$, we find as $\alpha_{1\rho}a \to 0$

$$\frac{H_0^{(1)}(i\alpha_{1\rho}a)}{H_1^{(1)}(i\alpha_{1\rho}a)} \sim -i\alpha_{1\rho}a\ln(i\alpha_{1\rho}a) \tag{72}$$

and (69) becomes

$$\lim_{\alpha_{1\rho}a\to 0}\left[\frac{\mu\epsilon_1 + \mu_1\epsilon}{k_\rho a}\frac{J_1'(k_\rho a)}{J_1(k_\rho a)} + \frac{\mu_1\epsilon_1 + \mu\epsilon}{k_\rho^2 a^2} + 2\mu\epsilon\ln(i\alpha_{1\rho}a)\right] = 0 \tag{73}$$

Thus the cutoff frequencies are determined from

$$J_1\left(k_{c1p}a\sqrt{1 - \frac{n^2}{n_1^2}}\right) = 0 \tag{74}$$

which gives for HE_{1p} modes

$$k_{c1p}a\sqrt{1 - \frac{n^2}{n_1^2}} = \xi_{1p} \tag{75}$$

where ξ_{1p} is the pth root of the first-order Bessel function $J_1(\xi)$. It is important to observe that the first root of J_1 is zero, implying that the cutoff frequency for the HE_{11} modes is zero. This is similar to the fundamental TE_{0p} and TM_{0p} modes with zero cutoff wavenumbers for symmetric slab waveguides. We can appreciate that as $\mu \to \mu_1$ and $\epsilon \to \epsilon_1$, the HE_{11} mode approaches a TEM wave for which there is no cutoff.

In the following, we list $k_c a(1 - n_1^2/n^2)^{1/2}$ in increasing order for the various guided modes by assuming $\mu\epsilon/\mu_1\epsilon_1 = 1.1$. Note that (i) TE_{0p} and TM_{0p} are degenerate modes in the sense that they have the same cutoff frequencies determined from (66); (ii) HE_{1p} and $\mathrm{EH}_{1(p-1)}$

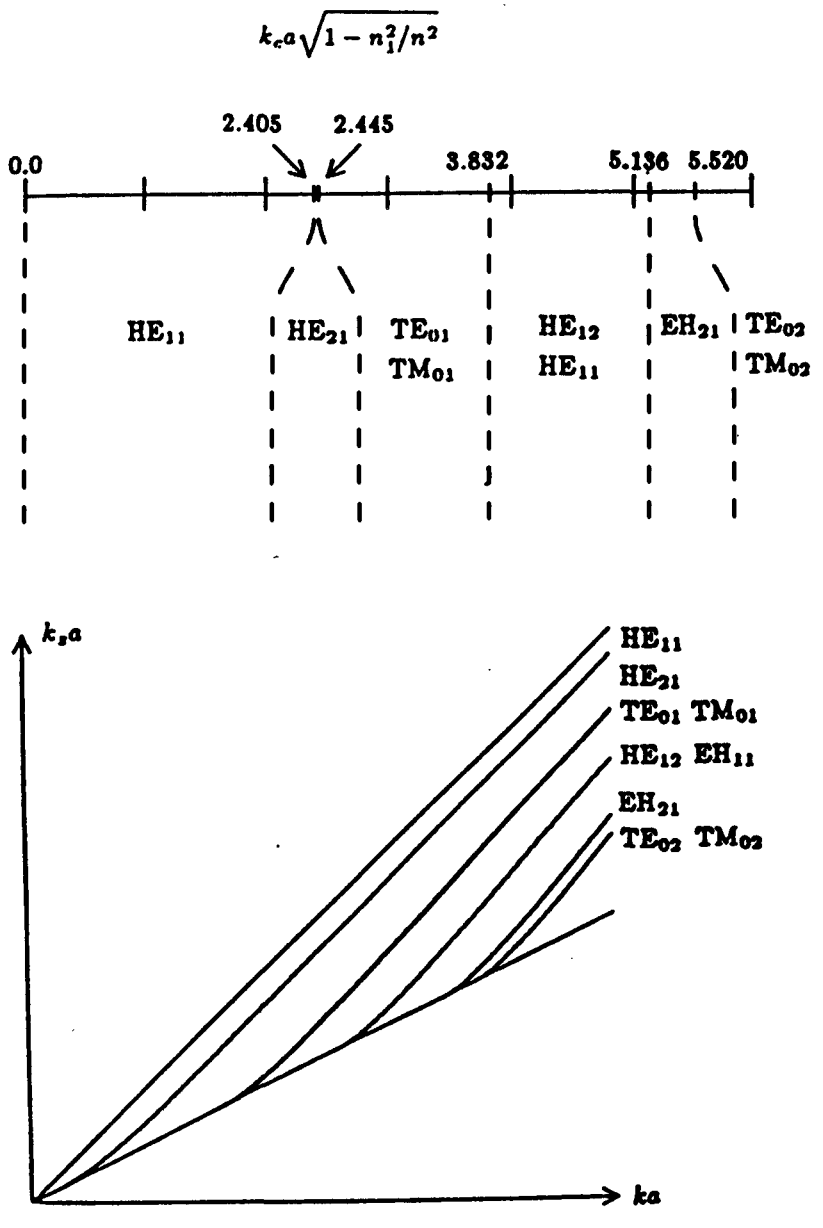


Figure 3.6.11 Cutoff of circular dielectric waveguide modes.

modes share the same cutoff frequencies because they are both determined from the roots of the first-order Bessel function $J_1(\xi)$ as seen from (75) and (68), except that for HE modes the first root is taken to be zero while for EH modes the first root is taken to be 3.832; (iii) the HE_{11} frequency operating range is very wide if the radius a is taken to be small and the refractive index n is only slightly larger than that of its surrounding medium n_1. For instance, in the case of a single-mode fiber waveguide for which $a \approx 1\mu m$, $n \approx 1.05$, and $1 - n_1^2/n^2 \approx 0.09$, the single-mode operating range extends from zero to $8 \times 10^6\ m^{-1}$.

3.7 Cavity Resonators

a. Rectangular Cavity Resonator

A resonator with uniform cross section in the $\hat{z}$ direction can be viewed as a waveguide with both ends closed. Instead of guided waves propagating along the z axis, the waves are standing in the $\hat{z}$ direction. The standing wave can be viewed as a superposition of a guided wave in the $+\hat{z}$ direction and a guided wave in the $-\hat{z}$ direction. The formulation for waveguides is also applicable to resonators. We have

$$\overline{E}_s = \frac{1}{k^2 - k_z^2}\left[\nabla_s \frac{\partial}{\partial z} E_z + i\omega\mu \nabla_s \times \overline{H}_z\right] \tag{1a}$$

$$\overline{H}_s = \frac{1}{k^2 - k_z^2}\left[\nabla_s \frac{\partial}{\partial z} H_z - i\omega\epsilon \nabla_s \times \overline{E}_z\right] \tag{1b}$$

and

$$\left(\nabla^2 + k^2\right) E_z = 0 \tag{2a}$$

$$\left(\nabla^2 + k^2\right) H_z = 0 \tag{2b}$$

where $k^2 = \omega^2\mu\epsilon$ and the Laplacian operation ∇^2 in (2) is now a three-dimensional operator.

Consider a metallic rectangular cavity as shown in Figure 3.7.1. It is a waveguide closed with metallic walls at $z = 0$ and $z = d$. To satisfy the boundary conditions, we find for TM modes,

$$E_z = E_{mnp} \sin\frac{m\pi x}{a} \sin\frac{n\pi y}{b} \cos\frac{p\pi z}{d} \tag{3a}$$

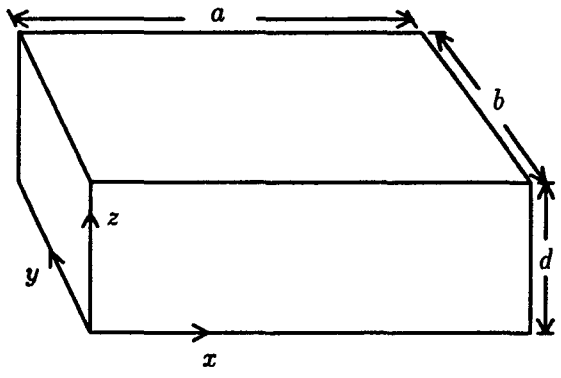


Figure 3.7.1 Rectangular cavity.

$$E_x = -\frac{E_{mnp}}{(m\pi/a)^2 + (n\pi/b)^2} \frac{m\pi}{a} \frac{p\pi}{d} \cos\frac{m\pi x}{a} \sin\frac{n\pi y}{b} \sin\frac{p\pi z}{d} \tag{3b}$$

$$E_y = -\frac{E_{mnp}}{(m\pi/a)^2 + (n\pi/b)^2} \frac{n\pi}{b} \frac{p\pi}{d} \sin\frac{m\pi x}{a} \cos\frac{n\pi y}{b} \sin\frac{p\pi z}{d} \tag{3c}$$

$$H_x = -\frac{i\omega\epsilon E_{mnp}}{(m\pi/a)^2 + (n\pi/b)^2} \frac{n\pi}{b} \sin\frac{m\pi x}{a} \cos\frac{n\pi y}{b} \cos\frac{p\pi z}{d} \tag{3d}$$

$$H_y = \frac{i\omega\epsilon E_{mnp}}{(m\pi/a)^2 + (n\pi/b)^2} \frac{m\pi}{a} \cos\frac{m\pi x}{a} \sin\frac{n\pi y}{b} \cos\frac{p\pi z}{d} \tag{3e}$$

and for TE modes

$$H_z = H_{mnp} \cos\frac{m\pi x}{a} \cos\frac{n\pi y}{b} \sin\frac{p\pi z}{d} \tag{4a}$$

$$H_x = -\frac{H_{mnp}}{(m\pi/a)^2 + (n\pi/b)^2} \frac{m\pi}{a} \frac{p\pi}{d} \sin\frac{m\pi x}{a} \cos\frac{n\pi y}{b} \cos\frac{p\pi z}{d} \tag{4b}$$

$$H_y = -\frac{H_{mnp}}{(m\pi/a)^2 + (n\pi/b)^2} \frac{n\pi}{b} \frac{p\pi}{d} \cos\frac{m\pi x}{a} \sin\frac{n\pi y}{b} \cos\frac{p\pi z}{d} \tag{4c}$$

$$E_x = -\frac{i\omega\mu H_{mnp}}{(m\pi/a)^2 + (n\pi/b)^2} \frac{n\pi}{b} \cos\frac{m\pi x}{a} \sin\frac{n\pi y}{b} \sin\frac{p\pi z}{d} \tag{4d}$$

$$E_x = \frac{i\omega\mu H_{mnp}}{(m\pi/a)^2 + (n\pi/b)^2} \frac{m\pi}{a} \sin\frac{m\pi x}{a} \cos\frac{n\pi y}{b} \sin\frac{p\pi z}{d} \tag{4e}$$

Substituting (3*a*) and (4*a*) into (2), we obtain the same dispersion relation for the TM and TE modes

$$k_r^2 = (m\pi/a)^2 + (n\pi/b)^2 + (p\pi/d)^2$$

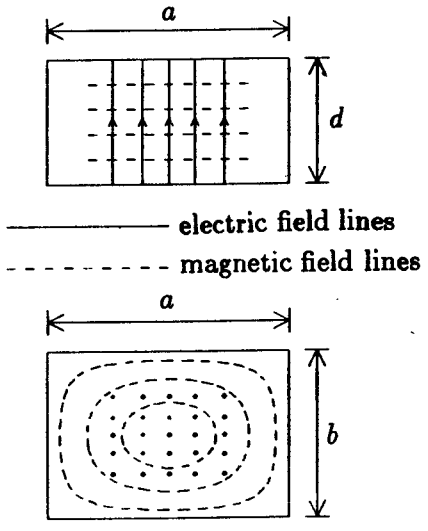


Figure 3.7.2 Field lines for TE_{101} mode.

This gives the resonant wavenumber $k_r = \omega_r(\mu\epsilon)^{1/2}$. The resonant wavenumbers for TM_{mnp} modes and TE_{mnp} modes are identical. It is interesting to observe that TM_{mn0} modes correspond to waveguide modes at cutoff, where $k_z = 0$.

When the resonator dimensions are such that $a > b > d$, the lowest resonant wavenumber is found to be

$$k_r = \sqrt{(\pi/a)^2 + (\pi/b)^2} \tag{5}$$

with $m = n = 1$ and $p = 0$. The mode inside the resonator is TM_{110}. The field distribution is illustrated in Figure 3.7.2. We see that the electric fields are perpendicular to the plate boundaries at $z = 0$ and $z = d$ and concentrate at the center of the cavity so that the tangential $\overline{E}$ field vanishes at the boundaries $x = 0$ and a and $y = 0$ and b. This field can also be viewed as a dominant waveguide mode propagating in the $\hat{y}$ direction and reflected at the walls $y = 0$ and $y = b$ to form a standing wave. If the labels of the coordinate axes y and z are interchanged, this mode may also be called a TE_{101} mode.

In all cavity resonators, the quality factor Q expresses the ratio of energy storage to energy dissipation. Let U be the energy stored in a resonator and P_d be the power dissipation in the resonator. We

define

$$Q = \frac{\omega_0 U}{P_d} \tag{6}$$

where ω_0 is the resonant angular frequency. Under the assumption of a lossless medium, we calculate

$$\begin{aligned} U &= \frac{1}{2} Re\left\{ \int_0^d dz \int_0^b dy \int_0^a dx \left[\frac{\epsilon}{2} |E|^2 + \frac{\mu}{2} |H|^2 \right] \right\} \\ &= \epsilon \frac{abd}{8} E_{110}^2 \end{aligned} \tag{7}$$

for the dominant TM_{110} mode in the rectangular cavity. Integrating over the cavity walls, we obtain

$$\begin{aligned} P_d =& \frac{1}{2} \sqrt{\omega_0 \mu / 2\sigma}\, Re \left\{ 2 \int_0^d dz \int_0^a dx\, |H_x|_{y=0}^2 \right. \\ &\left. + 2 \int_0^d dz \int_0^b dy\, |H_y|_{x=0}^2 + 2 \int_0^a dx \int_0^b dy \left(|H_x|^2 + |H_y|^2 \right)_{z=0} \right\} \\ =& \frac{1}{2} \sqrt{\omega_0 \mu / 2\sigma} \left[\frac{ad}{b^2} + \frac{bd}{a^2} + \frac{1}{2} \left(\frac{b}{a} + \frac{a}{b} \right) \right] \frac{\pi^2 \omega_0^2 \epsilon^2}{(\pi^2/a^2 + \pi^2/b^2)^2} E_{110}^2 \end{aligned}$$

Therefore,

$$Q = \sqrt{\frac{2\sigma}{\omega_0 \epsilon}}\, \frac{\pi d (a^2 + b^2)^{3/2}}{2[ab(a^2 + b^2) + 2d(a^3 + b^3)]} \tag{8}$$

In this derivation, we used the fact that $\omega_0 \sqrt{\mu\epsilon} = \sqrt{(\pi^2/a^2) + (\pi^2/b^2)}$. For a cubic cavity with $a = b = d = 2\,\mathrm{cm}$, the resonant frequency, according to (7), is 10 GHz; and the quality factor is $Q \approx 10^4$ when the cavity is air-filled and is made of copper walls. Other sources of loss, such as the material filling the cavity, surface irregularities of the cavity walls, and coupling with external systems, all contribute toward power dissipation P_d and thereby decrease Q.

b. Circular Cavity Resonator

Consider a circular cavity of height d and radius a [Fig. 3.7.3]. Under the assumption that $d < a$, the fundamental mode is TM_{010}, which corresponds to the waveguide mode TM_{01} at cutoff. The fields inside the cavity are

$$E_z = E_0 J_0(k\rho) \tag{9a}$$

$$H_\phi = -i \sqrt{\epsilon/\mu}\, E_0 J_1(k\rho) \tag{9b}$$

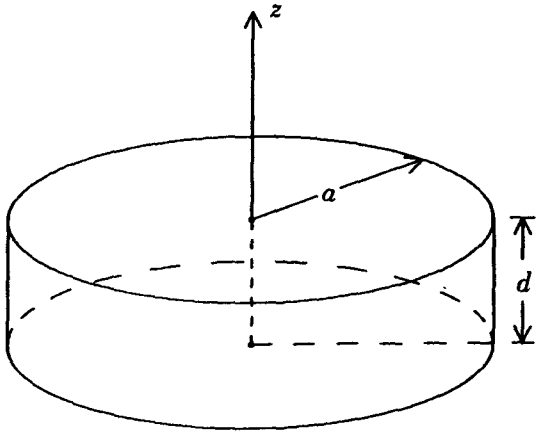


Figure 3.7.3 Circular cavity resonator.

The resonant wavenumber is

$$k_r a = 2.405 \tag{9c}$$

The time-average energy stored in the cavity is calculated as

$$\begin{aligned} U &= \frac{1}{2}\int_0^a 2\pi\rho\, d\rho \left[\frac{\epsilon}{2}|E_z|^2 + \frac{\mu}{2}|H_\phi|^2\right] d \\ &= E_0^2 \frac{\pi\epsilon d}{2} a^2 J_1^2(ka) \end{aligned} \tag{10}$$

The integral formula for Bessel functions,

$$\int \rho\, d\rho\, B_m^2(k\rho) = \frac{\rho^2}{2}\left[B_m'^2(k\rho) + \left(1 - \frac{m^2}{k^2\rho^2}\right) B_m^2(k\rho)\right] \tag{11}$$

is used, and so is $J_0(ka) = 0$. The power dissipation caused by wall loss is

$$\begin{aligned} P_d &= \frac{E_0^2}{2}\sqrt{\omega_0\mu/2\sigma}\left[2\pi a d \frac{\epsilon}{\mu} J_1^2(ka) + 2\int_0^a 2\pi\rho \frac{\epsilon}{\mu} J_1^2(k\rho)\, d\rho\right] \\ &= \sqrt{\omega_0\mu/2\sigma}\, E_0^2 \frac{\epsilon}{\mu}\pi a(d+a) J_1^2(ka) \end{aligned} \tag{12}$$

The first term is due to loss on the side wall and the second term to loss on the walls at $z = 0$ and $z = d$. The quality factor becomes

$$Q = \frac{\omega_0 U}{P_d} = \sqrt{2\sigma/\omega_0\epsilon}\, \frac{2.405}{2(1 + a/d)} \tag{13}$$

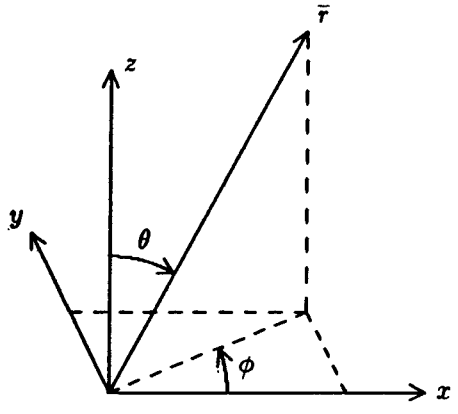


Figure 3.7.4 Spherical coordinate system.

where we made use of (9c). In the mode designation TM_{010} the three subscripts correspond to ϕ, ρ, and z variations, respectively . The TE_{011} mode, for instance, is the waveguide mode TE_{01} forming a standing wave in the $\hat{z}$ direction.

c. Spherical Cavity Resonator

For a spherical cavity, the waveguide formulation breaks down because there is no uniform cross section in any direction. Consider Maxwell's equations in spherical coordinates [Fig. 3.7.4], and treat the case with ϕ symmetry, $\partial/\partial\phi = 0$. Instead of decomposing a general field into TM and TE to $\hat{z}$ components, we decompose into TM and TE to $\hat{r}$ components. For the TM waves, Maxwell's equations give

$$\frac{\partial}{\partial r}(rE_\theta) - \frac{\partial E_r}{\partial \theta} = i\omega\mu r H_\phi \tag{14a}$$

$$\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(H_\phi \sin\theta) = -i\omega\epsilon E_r \tag{14b}$$

$$-\frac{\partial}{\partial r}(rH_\phi) = -i\omega\epsilon r E_\theta \tag{14c}$$

Inserting (14b) and (14c) into (14a) yields an equation for H_ϕ

$$\frac{1}{r}\frac{\partial^2}{\partial r^2}(rH_\phi) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial H_\phi}{\partial\theta}\right) - \frac{1}{r^2\sin^2\theta}H_\phi + k^2 H_\phi = 0 \tag{15}$$

A similar equation for E_ϕ, which is the dual of (15), can be obtained for the TE waves.

Before solving for (14) and (15), we first study general solutions to Helmholtz wave equations in spherical coordinates. From source-free Maxwell's equations in isotropic media, the wave equation for $\overline{E}$ and $\overline{H}$ is readily derived

$$(\nabla^2 + k^2)\left\{\begin{matrix}\overline{E}\\ \overline{H}\end{matrix}\right\} = 0$$

Let $W(r,\theta,\phi)$ denote any rectangular component of $\overline{E}$ or $\overline{H}$. The Helmholtz equation in spherical coordinates takes the form

$$\frac{1}{r}\frac{\partial^2}{\partial r^2}(rW) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial W}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 W}{\partial\phi^2} + k^2 W = 0 \quad (16)$$

Equation (15) can be derived directly from (16) by noting that for $\overline{H} = \hat{\phi}H_\phi$, $\nabla^2\hat{\phi} = -\hat{\phi}\,1/(r^2\sin^2\theta)$. The solution to the Helmholtz equation (16) is obtained by separation of variables

$$W = R(r)\Theta(\theta)\Phi(\phi) \quad (17)$$

The special functions satisfy the following differential equations:

$$r\frac{d^2}{dr^2}(rR) + [(kr)^2 - n(n+1)]R \mp 0 \quad (18a)$$

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right) + \left[n(n+1) - \frac{m^2}{\sin^2\theta}\right]\Theta = 0 \quad (18b)$$

$$\frac{d^2\Phi}{d\phi^2} + m^2\Phi = 0 \quad (18c)$$

Solutions to (18a), (18b), and (18c) are, respectively, the spherical Bessel functions, $b_n(kr)$, the associated Legendre polynomials, $L_n^m(\cos\theta)$, and the harmonic functions, $e^{\pm im\phi}$.

The spherical Bessel functions $b_n(\xi)$ are related to the cylindrical Bessel functions $B_n(\xi)$ which satisfy the Bessel equation

$$\frac{d^2}{d\xi^2}B(\xi) + \frac{1}{\xi}\frac{d}{d\xi}B(\xi) + \left[1 - \frac{(n+1/2)^2}{\xi^2}\right]B(\xi) = 0$$

We can cast (18a) in the form of the Bessel equation by letting $R(\xi) = (\pi/2\xi)^{1/2}B(\xi)$ with $\xi = kr$. We thus find that

$$b_n(kr) = \sqrt{\pi/2kr}\, B_{n+1/2}(kr) \tag{19}$$

If n is an integer, $B_{n+1/2}$ reduces to simple sinusoids and powers of r. For the first few orders, for instance

$$j_0(kr) = \frac{\sin kr}{kr} \tag{20a}$$

$$j_1(kr) = -\frac{\cos kr}{kr} + \frac{\sin kr}{(kr)^2} \tag{20b}$$

$$j_2(kr) = -\frac{\sin kr}{kr} - \frac{3\cos kr}{(kr)^2} + \frac{3\sin kr}{(kr)^3} \tag{20c}$$

$$n_0(kr) = -\frac{\cos kr}{kr} \tag{21a}$$

$$n_1(kr) = -\frac{\sin kr}{kr} - \frac{\cos kr}{(kr)^2} \tag{21b}$$

$$n_2(kr) = \frac{\cos \sin kr}{kr} - \frac{3\sin kr}{(kr)^2} - \frac{3\cos kr}{(kr)^3} \tag{21c}$$

The spherical Hankel functions of the first kind take the form

$$h_0^{(1)}(kr) = \frac{e^{ikr}}{ikr} \tag{22a}$$

$$h_1^{(1)}(kr) = -\frac{e^{ikr}}{kr}\left(1 + \frac{i}{kr}\right) \tag{22b}$$

$$h_2^{(1)}(kr) = \frac{ie^{ikr}}{kr}\left[1 + \frac{3i}{kr} + 3\left(\frac{i}{kr}\right)^2\right] \tag{22c}$$

The spherical Hankel function of the second kind is the complex conjugate of $h_n^{(1)}$.

The first few orders of the associated Legendre polynomials of degree 1 take these forms:

$$P_0^1(\cos\theta) = 0 \tag{23a}$$

$$P_1^1(\cos\theta) = \sin\theta \tag{23b}$$

$$P_2^1(\cos\theta) = 3\sin\theta\cos\theta \tag{23c}$$

It is a general property that all of the associated Legendre polynomials $P_n^1(\cos\theta)$ are zero at $\theta = 0$ and π; at $\theta = \pi/2$, they are zero if n is even and maximum if n is odd. For the H_ϕ component,

$$\overline{H} = \hat{\phi} H_\phi = (-\hat{x}\sin\phi + \hat{y}\cos\phi) H_\phi$$

Substituting in (16) yields

$$\left(\nabla^2 + k^2 - \frac{1}{r^2 \sin^2\theta}\right) H_\phi = 0$$

The effect of the last term on the solution is to increase the associated Legendre polynomial by one more degree in m.

In view of (16) and its solution in (17), we see that the solutions for H_ϕ in (15) take the form

$$H_\phi = b_n(kr) P_n^1(\cos\theta) \tag{24}$$

There is no ϕ dependence. For a spherical cavity with radius a, the spherical Bessel function is used because the origin is included. For the lowest TM mode we let $n = 1$ and use three subscripts on TM to denote the variations around r, ϕ, and θ, respectively. The TM_{101} mode has the field solutions

$$H_\phi = H_0 \sin\theta \sqrt{\frac{\pi}{2kr}}\, J_{3/2}(kr)$$

$$= H_0 \frac{\sin\theta}{kr}\left(\frac{\sin kr}{kr} - \cos kr\right) \tag{25a}$$

$$E_r = i2H_0 \sqrt{\frac{\mu}{\epsilon}}\, \frac{\cos\theta}{k^2 r^2}\left(\frac{\sin kr}{kr} - \cos kr\right) \tag{25b}$$

$$E_\theta = -iH_0 \sqrt{\frac{\mu}{\epsilon}}\, \frac{\sin\theta}{k^2 r^2}\left(\frac{k^2 r^2 - 1}{kr}\sin kr + \cos kr\right) \tag{25c}$$

The boundary condition of vanishing E_θ at $r = a$ gives

$$\tan ka = \frac{ka}{1 - k^2 a^2}$$

Solving this transcendental equation yields $ka \approx 2.74$, which gives the resonant wavenumber of the cavity.

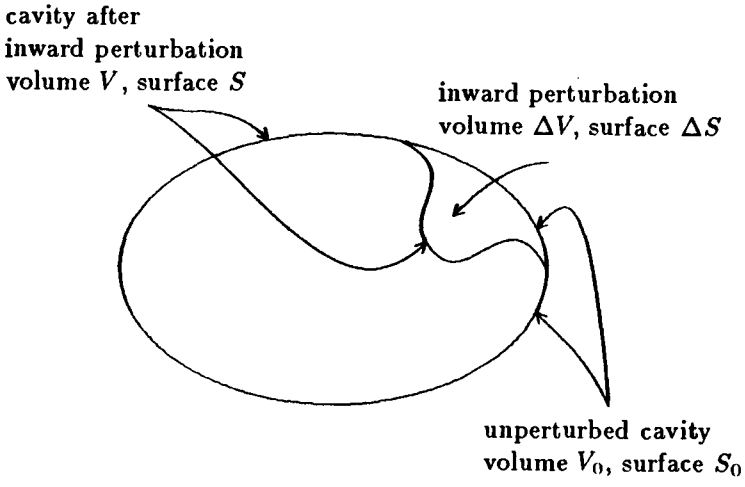


Figure 3.7.5 Cavity perturbation.

d. Cavity Perturbation

The resonant frequency of a cavity changes when a small perturbation is applied to either the cavity wall or the medium inside the cavity. First, we consider an inward perturbation of the cavity wall [Fig. 3.7.5]. The unperturbed fields have resonant frequency ω_0 and satisfy Maxwell's equations:

$$\nabla \times \overline{E}_0 = i\omega_0 \mu \overline{H}_0 \tag{26a}$$

$$\nabla \times \overline{H}_0 = -i\omega_0 \epsilon \overline{E}_0 \tag{26b}$$

With the perturbation, the resonant frequency becomes ω and the fields satisfy Maxwell's equations:

$$\nabla \times \overline{E} = i\omega \mu \overline{H} \tag{27a}$$

$$\nabla \times \overline{H} = -i\omega \epsilon \overline{E} \tag{27b}$$

The task is to calculate the deviation of ω from ω_0. We dot-multiply the complex conjugate of (26*a*) by $\overline{H}$ and subtract (27*b*), dot-multiplied by $\overline{E}_0^*$. The result is

$$\nabla \cdot \left(\overline{E}_0^* \times \overline{H}\right) = -i\omega_0 \mu \overline{H} \cdot \overline{H}_0^* + i\omega \epsilon \overline{E} \cdot \overline{E}_0^* \tag{28a}$$

Next, we dot-multiply (27*a*) by $\overline{H}_0^*$ and subtract the complex conjugate of (26*b*) dot-multiplied by $\overline{E}$. The result is

$$\nabla \cdot \left(\overline{E} \times \overline{H}_0^*\right) = i\omega \mu \overline{H} \cdot \overline{H}_0^* - i\omega_0 \epsilon \overline{E} \cdot \overline{E}_0^* \tag{28b}$$

Integrating the sum of (28a) and (28b) over the unperturbed volume $V_0 = V + \Delta V$, we find

$$\oint\!\!\!\!\!\!\int_{\Delta S} d\overline{S} \cdot \overline{E} \times \overline{H}_0^* = i(\omega - \omega_0) \iiint_{V_0} dV \, (\epsilon \overline{E} \cdot \overline{E}_0^* + \mu \overline{H} \cdot \overline{H}_0^*)$$

In these calculations, we use the fact that tangential $\overline{E}$ vanishes on the perturbed cavity surface and tangential $\overline{E}_0$ vanishes on the unperturbed cavity surface. Note that the surface integration extends over the small perturbed surface ΔS, while the volume integration extends over the unperturbed volume. We obtain the exact equation

$$\omega - \omega_0 = -i \frac{\oint\!\!\!\!\!\!\int_{\Delta S} d\overline{S} \cdot \overline{E} \times \overline{H}_0^*}{\iiint_{V_0} dV \, (\epsilon \overline{E} \cdot \overline{E}_0^* + \mu \overline{H} \cdot \overline{H}_0^*)} \tag{29}$$

Now we assume that the perturbation is so small that we can replace $\overline{E}$ and $\overline{H}$ on the right-hand side of (29) by their unperturbed values $\overline{E}_0$ and $\overline{H}_0$ to obtain approximate values for $\omega - \omega_0$:

$$\begin{aligned} \omega - \omega_0 &\approx -i \frac{\oint\!\!\!\!\!\!\int_{\Delta S} d\overline{S} \cdot \overline{E}_0 \times \overline{H}_0^*}{\iiint_{V_0} dV \left(\epsilon \left| \overline{E}_0 \right|^2 + \mu \left| \overline{H}_0 \right|^2 \right)} \\ &= \omega_0 \frac{\iiint_{\Delta V} dV \left(\mu \left| \overline{H}_0 \right|^2 - \epsilon \left| \overline{E}_0 \right|^2 \right)}{\iiint_{V_0} dV \left(\mu \left| \overline{H}_0 \right|^2 + \epsilon \left| \overline{E}_0 \right|^2 \right)} \\ &= \omega_0 \frac{\Delta W_m - \Delta W_e}{W_m + W_e} \end{aligned} \tag{30}$$

The denominator is the unperturbed total energy stored in the cavity. The numerator is the difference between the magnetic energy and the electric energy removed by the inward perturbation. Thus, if the inward perturbation is made at a place of large magnetic field, the resonant frequency is raised; if it is made at a place of large electric field, the resonant frequency is lowered. An opposite effect occurs for an outward perturbation.

Next, we investigate the resonant frequency change caused by material perturbation inside the cavity. Let the unperturbed medium be isotropic. To be more general, we include anisotropy in the perturbation. Maxwell's equations before and after perturbation are

$$\nabla \times \overline{E}_0 = i\omega_0 \mu \overline{H}_0 \tag{31a}$$

$$\nabla \times \overline{H}_0 = -i\omega_0 \epsilon \overline{E}_0 \tag{31b}$$

and

$$\nabla \times \overline{E} = i\omega\mu\overline{H} + i\omega\Delta\overline{\overline{\mu}} \cdot \overline{H} \tag{32a}$$

$$\nabla \times \overline{H} = -i\omega\epsilon\overline{E} - i\omega\Delta\overline{\overline{\epsilon}} \cdot \overline{E} \tag{32b}$$

We dot-multiply the complex conjugate of (31a) by $\overline{H}$ and subtract (32b) dot-multiplied by $\overline{E}_0^*$. The result is

$$\nabla \cdot (\overline{E}_0^* \times \overline{H}) = -i\omega_0\mu\overline{H}_0^* \cdot \overline{H} + i\omega\epsilon\overline{E} \cdot \overline{E}_0^* + i\omega(\Delta\overline{\overline{\epsilon}} \cdot \overline{E}) \cdot \overline{E}_0^*$$

A similar operation on (31b) and (32a) gives

$$\nabla \cdot (\overline{E} \times \overline{H}_0^*) = i\omega\mu\overline{H} \cdot \overline{H}_0^* + i\omega(\Delta\overline{\overline{\mu}} \cdot \overline{H}) \cdot \overline{H}_0^* - i\omega_0\epsilon\overline{E} \cdot \overline{E}_0^*$$

Integrating the sum of these two equations over the cavity volume and making use of the boundary condition that both $\hat{n} \times \overline{E} = 0$ and $\hat{n} \times \overline{E}_0 = 0$ on the cavity surface, we obtain

$$\frac{\omega - \omega_0}{\omega} = \frac{-\iiint_V dV \left[(\Delta\overline{\overline{\mu}} \cdot \overline{H}) \cdot \overline{H}_0^* + (\Delta\overline{\overline{\epsilon}} \cdot \overline{E}) \cdot \overline{E}_0^*\right]}{\iiint_V dV (\mu\overline{H} \cdot \overline{H}_0^* + \epsilon\overline{E} \cdot \overline{E}_0^*)} \tag{33}$$

This is also an exact formula. When the perturbation is so small that we can replace the perturbed fields on the right-hand side by their unperturbed values, we obtain

$$\frac{\omega - \omega_0}{\omega} \approx -\frac{\Delta W_m + \Delta W_e}{W_m + W_e} \tag{34}$$

The denominator expresses the unperturbed total energy inside the cavity, and the numerator corresponds to the increase in magnetic and

electric energies caused by the material perturbation. Thus any increase in the permeability or the permittivity of the material inside the cavity decreases the resonant frequency. For instance, recall that the resonant wavenumber for the dominant mode in a circular resonator is $k_r a = 2.405$. As $k_r = \omega\sqrt{\mu\epsilon}$ increases, the resonant frequency ω_0 decreases. The material, of course, need not be uniformly distributed throughout the cavity. The calculation of ΔW_m and ΔW_e corresponding to $\Delta\overline{\overline{\mu}}$ and $\Delta\overline{\overline{\epsilon}}$ extends only over the region where perturbation occurs.

PROBLEMS

Problem P3.1

When the incident k vector is normal to a plane boundary, a TE wave becomes a TEM wave; a TM wave also becomes a TEM wave. Compare the reflection and transmission coefficients for TE and TM waves at normal incidence. Do both TE and TM results reduce to the same number? If not, why? Do the reflectivities and transmissivities for TE and TM waves at normal incidence reduce to the same result?

Problem P3.2

For a highly conducting earth, assume $\sigma/\omega\epsilon \gg 1$, where σ is the conductivity of the earth. Consider a TM wave incident upon the boundary of the earth surface.

(a) Write the magnetic and electric field components for $R^{\mathrm{TM}} \neq 0$ and then set $R^{\mathrm{TM}} = 0$. Show that

$$k_z = k_{zR} + ik_{zI} = \frac{\omega\epsilon}{\sigma}\sqrt{\frac{\omega\mu\sigma}{2}}(1-i)$$

which propagates towards the surface and attenuates away from the surface. This is known as the Zenneck wave. Show that the Zenneck wave is a fast wave with $Re\,(k_z) < k$.

(b) Define a surface impedance

$$R_s - iX_s = k_z/\omega\epsilon$$

Show that for the conducting earth

$$R_s = X_s = \sqrt{\frac{\omega\mu}{2\sigma}}$$

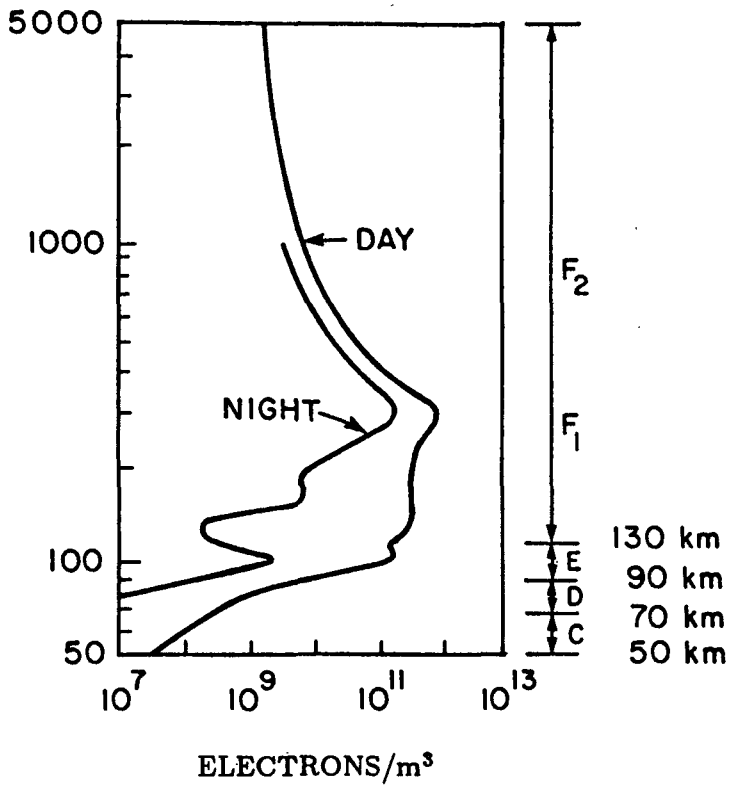


Figure P3.3 Electron density profiles(typical).

(c) For an inductive surface characterized by $R_s = 0$, write down $\overline{H}$ with $R^{TM} = 0$ and discuss the behavior of the resulting wave.

Problem P3.3

The ionosphere extends from approximately 50 km above the earth to several earth radii (mean earth radius is about 6371 km) with the maximum in ionization density at about 300 km. The ionization density profile shows ledges where it varies more slowly with altitude. These ledges are the C, D, E, F_1, and F_2 layers [Fig. P3.3]. These maxima arise because both the solar radiation and the composition of the atmosphere change with altitude. The heights and the intensities of ionization of these layers change with the hour of the day, the season of the year, the sunspot cycle, etc. The electron density varies from approximately $10^7/m^3$ to $10^{12}/m^3$ in going from the lowest to

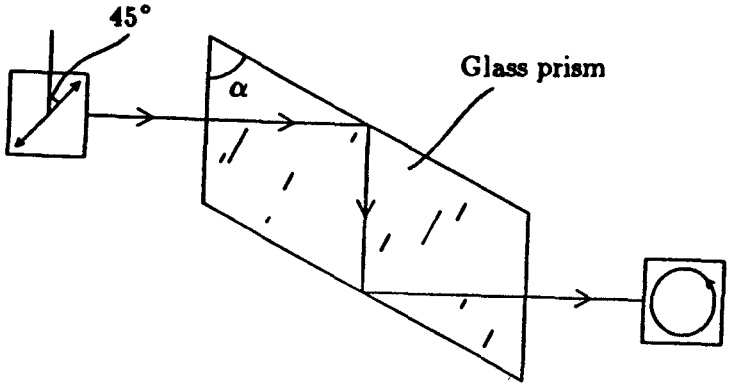


Figure P3.5

the highest layer. For simplicity, assume that the ionosphere consists of a 40 km thick E-layer with electron density $N = 10^{11}/m^3$ below a 200 km thick F-layer with $N = 6 \times 10^{11}/m^3$.

(a) What are the plasma frequencies of the E and F layers ?

(b) Consider a plane wave of 10 MHz incident at an angle θ upon the ionosphere from below the E-layer, what is the angle θ_t of the ionosphere wave?

(c) Let $\theta = 30°$, below what frequency will the wave be totally reflected by the E-layer and below what frequency will it be totally reflected by the F-layer ?

Problem P3.4

A plane wave in free space is incident at an angle θ on a conducting half-space. Show that for large $\sigma/\omega\epsilon_o$ the transmitted angle is

$$\theta_t \approx \tan^{-1}\left[\sqrt{\frac{2\omega\epsilon_o}{\sigma}}\sin\theta\right]$$

which is a very small angle, and the transmitted wave is almost perpendicular to the boundary.

Problem P3.5

Show that the difference in relative phase change between TM and TE waves upon total internal reflection is

$$\Delta = \phi_{\mathrm{TE}} - \phi_{\mathrm{TM}} = 2\tan^{-1}\left[\frac{\cos\theta\left(\sqrt{\sin^2\theta - n^{-2}}\right)}{\sin^2\theta}\right]$$

The phase change produced on internal reflection may be utilized to obtain circularly polarized light from linearly polarized light. The scheme, devised by Fresnel, is shown in Figure P3.5. The essential element is a glass prism ($n = 1.6$) made in the form of a rhomb having an apex angle α. Linearly polarized light with a direction of polarization at an angle of 45° with respect to the face edge of the rhomb enters normally on one face. What should α be so that the light coming out is circularly polarized?

Problem P3.6

Let a plane wave be incident on a plane boundary from the inside of a negative uniaxial crystal. Consider the special case in which the optic axis is perpendicular to the plane of incidence. Find the range of θ such that there is total internal reflection for the extraordinary wave.

Problem P3.7

A plane wave is incident upon a half-space medium with $\epsilon = 10\epsilon_o$, $\mu = \mu_o$, and $\sigma = 10^{-4}$ mhos/m. The transmitted wave is picked up by a receiver located at a distance $d = 100$ m from the interface. At what frequency will the receiver pick up a maximum transmitted power?

Problem P3.8

It is interesting to compare the following two cases: (i) at total reflection, the incident power is totally reflected; and (ii) when the transmitting medium is a perfect conductor, the incident power is also totally reflected. As far as the reflected TE wave is concerned, it makes no difference whether the wave is reflected by a perfect conductor or by a dielectric boundary if both boundaries render the reflected wave with identical phases.

(a) Consider a perfectly conducting boundary at $x = -d$. Show that the wave is phase-shifted by $-\pi$ at the perfectly conducting boundary.

(b) The path length from $x = 0$ to $x = -d$ provides another phase shift, $2k_x d$, for the wave. Find d such that the reflected TE wave

at $x > 0$ experiences the same amount of phase shift in both cases.

Experiments by Goos and Hänchen have demonstrated that a beam of light is laterally shifted when totally reflected at a dielectric boundary. We must note, however, that although this argument can be used to display a lateral shift for the reflected $\bar{k}$ vector, the analysis is true only for plane waves, and there is no way to observe a lateral shift experimentally for plane waves.

Problem P3.9

Compare the phenomena of total reflection for $\theta > \theta_C$ and total transmission for $\theta = \theta_B$ at an isotropic dielectric interface.

(a) Total reflection occurs at a range of incident angles larger than the critical angle θ_C; total transmission of TM waves occurs only at the Brewster angle θ_B.

(b) Total reflection occurs only when the incident medium is denser than the transmitted medium. The Brewster angle occurs for any two media.

(c) When an unpolarized wave is totally reflected, the reflected wave is still unpolarized. When the TM wave components of an unpolarized wave are totally transmitted, the reflected wave contains only TE waves.

Suppose a TM wave is incident at an angle θ such that $\theta = \theta_B > \theta_C$. Then the wave is totally transmitted and at the same time it is totally reflected. Explain what is happening.

Problem P3.10

A Nicol prism made of calcite is cut diagonally and then joined together with a film of Canada balsam (refractive index $n = 1.53$). Calcite is a negative uniaxial crystal with $\sqrt{\epsilon_z/\epsilon} = 1.49/1.66$. Show that with the arrangement shown in Figure P3.10 an incident light from the left becomes a linearly polarized light when it leaves the crystal from the right.

Problem P3.11

For a cold plasma with collisions,

$$\epsilon(\omega) = \epsilon_o \left[1 - \frac{\omega_p^2}{\omega^2 + \nu^2} + \frac{i\nu\omega_p}{\omega(\omega^2 + \nu^2)} \right]$$

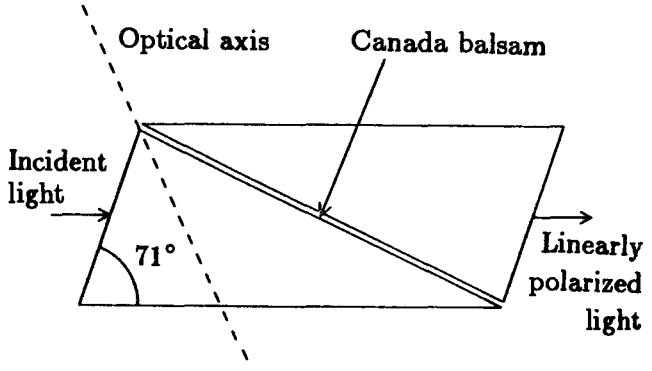


Figure P3.10

Consider a plane wave incident on this medium. Determine $\overline{k}_t$ in magnitude and direction as a function of ν.

Problem P3.12

Consider a solid-state Fabry-Perot etalon filter made of an eight-layer stratified medium. Regions 1, 3, 5, and 7 are made of magnesium fluoride (refractive index $n = 1.35$) and are a quarter-wavelength thick. Regions 2, 4, and 6 are made of zinc sulfide (refractive index $n = 2.3$). Regions 2 and 6 are a quarter-wavelength thick, but region 4 is a half-wavelength thick. What are the reflectivity and transmissivity for a plane wave normally incident upon this stratified medium? Explain why the structure can be used for filtering purposes.

Problem P3.13

A plane wave is totally reflected when incident upon a glass–air boundary. At this incident angle, another piece of glass is brought very close to the first one so that there is a very small air gap between the two. Calculate the reflection and transmission coefficients as a function of the gap dimension. Show that transmission is now possible.

Problem P3.14

Sun light glares caused by reflections from plane surfaces are partially linearly polarized.

(a) Plot the reflectivities for the TE and TM waves as functions of

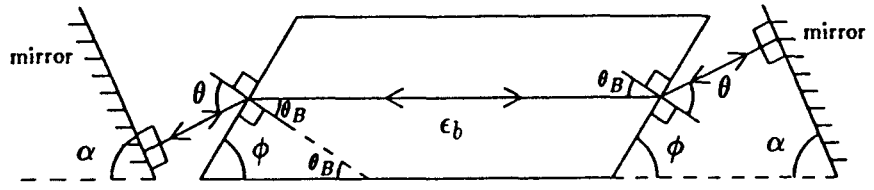


Figure P3.15

incident angle.

$$r^{\mathrm{TE}} = |R^{\mathrm{TE}}|^2 = \left| \frac{1 - k_t \cos\theta_t / k \cos\theta}{1 + k_t \cos\theta_t / k \cos\theta} \right|^2$$

$$r^{\mathrm{TM}} = |R^{\mathrm{TM}}|^2 = \left| \frac{1 - \epsilon_0 k_t \cos\theta_t / \epsilon_t k \cos\theta}{1 + \epsilon_0 k_t \cos\theta_t / \epsilon_t k \cos\theta} \right|^2$$

where by phase matching $k_t \sin\theta = k \sin\theta$. Show that for all incident angles, there are more reflected TE components than TM components $|R^{\mathrm{TE}}|^2 \geq |R^{\mathrm{TM}}|^2$.

(b) Determine the Brewster angle for $\epsilon_t = 9$. The Brewster angle, θ_B, is also called the polarization angle because at θ_B the reflected wave is entirely TE polarized.

(c) Your polaroid glasses absorb one linear component of incident light. To minimize sun glare, what component, TE or TM, reaches your eyes after passing through the glasses ? Explain why.

Problem P3.15

A gas laser is often a tube containing gas, fitted with Brewster-angle windows and external mirrors. The output of the laser will be linearly polarized. For what reason and in which direction? Let a solid-state laser be fabricated of rods with ends beveled at the Brewster angle θ_B [Fig. P3.15]. Sketch at what locations and at what angles you would place the external mirrors. Calculate and indicate in your sketch all appropriate angles, including the bevel angle of the glass rod, which has a dielectric constant $\epsilon = 2.5$.

Problem P3.16

In the microwave remote sensing of the Earth from satellite or aircraft, a radiometer is used to measure the emissivity of the area under observation. The emissivity e is related to reflectivity or to the power reflection coefficient by $e = 1 - r$. Theoretically, we should be able to determine, for instance, the ice thickness on a lake. Assume that the lake ice permittivity is $\epsilon = 3.2(1 + i0.01)\epsilon_o$ and that the water is a perfect reflector. Discuss the frequency range that you would recommend and the depth that the radiometer can "see" through the ice.

Problem P3.17

A plane wave linearly polarized along the $\hat{x}$ direction is propagating along the $\hat{z}$ direction and is normally incident upon an anisotropic plasma slab. Find the reflection and transmission coefficients when the dc magnetic field is along the (i) $\hat{z}$ direction, (ii) $\hat{y}$ direction, (iii) $\hat{x}$ direction.

Problem P3.18

A plane wave is incident from inside a positive uniaxial medium upon the interface with free space. The optic axis of the uniaxial medium is inclined at 45° with respect to the interface, and the incident angle of the wave is perpendicular to the optic axis. Find the reflection and transmission coefficients and the reflectivity and transmissivity.

Problem P3.19

Determine the propagation matrices for TM waves in a stratified media.

Problem P3.20

Using a parallel derivation of the forward-propagation matrix show that the backward-propagation matrix is

$$\overline{\overline{U}}_{l(l+1)} = \frac{1}{2}(1 + p_{l(l+1)}) \begin{pmatrix} e^{ik_{(l+1)z}(d_{l+1}-d_l)} & R_{l(l+1)}e^{-ik_{(l+1)z}(d_{l+1}-d_l)} \\ R_{l(l+1)}e^{ik_{(l+1)z}(d_{l+1}-d_l)} & e^{-ik_{(l+1)z}(d_{l+1}-d_l)} \end{pmatrix}$$

Show that $\overline{\overline{U}}_{l(l+1)} \cdot \overline{\overline{V}}_{(l+1)l} = \overline{\overline{V}}_{(l+1)l} \cdot \overline{\overline{U}}_{l(l+1)} = \overline{\overline{I}}$ where $\overline{\overline{I}}$ is the 2×2

identity matrix. Prove that the reflection and transmission coefficients can be calculated from

$$\begin{pmatrix} Re^{-ik_z d_0} \\ e^{ik_z d_0} \end{pmatrix} = \overline{\overline{U}}_{0t} \cdot \begin{pmatrix} 0 \\ T \end{pmatrix} = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix} \begin{pmatrix} 0 \\ T \end{pmatrix}$$

where

$$\overline{\overline{U}}_{0t} = \overline{\overline{U}}_{01} \cdot \overline{\overline{U}}_{12} \cdots \overline{\overline{U}}_{nt}$$

What is $\overline{\overline{U}}_{nt}$?

Problem P3.21

A plane wave of angular frequency ω is incident on a plasma medium with permeability μ_o and permittivity

$$\epsilon = \epsilon_o \left(1 - \frac{\omega_p^2}{\omega^2}\right)$$

where ω_p is the plasma frequency and $\omega = 2\omega_p$.

(a) Calculate the critical angle θ_C such that the incident wave is totally reflected.

(b) Calculate the Brewster angle θ_B such that TM waves are totally transmitted.

(c) In general for any two isotropic media, can you find an incident angle θ such that $\theta = \theta_B > \theta_C$? If you can, give an example. If you cannot, explain why not.

Problem P3.22

A ferrite medium occupies the half-space $x > 0$ and is magnetized along the z axis by a static magnetic field. Let a TE wave be incident from $x < 0$ at an angle θ_i. Find the reflection and the transmission coefficients. Show that the reflection coefficient of a TE wave incident from the $(-\theta_i)$ direction is the complex conjugate of the reflection coefficient for the first case. This is an example of nonreciprocal behavior. Here the permeability tensor is of the form

$$\mu = \begin{pmatrix} \mu & -i\mu_g & 0 \\ i\mu_g & \mu & 0 \\ 0 & 0 & \mu_o \end{pmatrix}$$

Problem P3.23

Show that for a three-layer medium, by letting $t = 3$ and $n = 2$ in 3.3.24, the reflection coefficient is

$$R = \Big\{ \Big[R_{01} + R_{12} e^{i2k_{1z}(d_1 - d_0)} \Big] + \Big[R_{01} R_{12} + e^{i2k_{1z}(d_1 - d_0)} \Big] R_{23} e^{i2k_{2z}(d_2 - d_1)} \Big\} e^{i2k_z d_0} \Big/ \Big\{ \Big[1 + R_{01} R_{12} e^{i2k_{1z}(d_1 - d_0)} \Big] + \Big[R_{12} + R_{01} e^{i2k_{1z}(d_1 - d_0)} \Big] R_{23} e^{i2k_{2z}(d_2 - d_1)} \Big\}$$

Show that this result reduces to (3.3.26) under the following limits: (i) $d_1 = d_0$ and $R_{01} = 0$; (ii) $d_1 = d_0$ or $d_1 = d_2$ and $R_{12} = 0$ which also implies $k_{1z} = k_{2z}$; and (iii) $R_{23} = 0$.

Problem P3.24

For an inhomogeneous medium with sufficiently small spatial variation, approximate solutions can be obtained using the WKB (Wentzel–Kramers–Brillouin) approach. For a half-space medium with a permittivity profile $\epsilon_t(z)$, we may write the solution for the TE waves as

$$E_{ty}(x, z) = E_t(z) e^{ik_x x}$$

(a) Show that the equation governing $E_t(z)$ is

$$\left[\frac{d^2}{dz^2} + k_{tz}^2(z) \right] E_t(z) = 0$$

where $k_{tz}^2 = k_t^2(z) - k_x^2$, and $k_t^2(z) = \omega^2 \mu \epsilon_t(z)$.

(b) The WKB approximation or the geometrical optics approximation to the second order differential equation takes the form

$$E_{tz} = \frac{A_1}{\sqrt{k_{tz}(z)}} e^{i \int_0^z dz' k_{tz}(z')} + \frac{A_2}{\sqrt{k_{tz}(z)}} e^{-i \int_0^z dz' k_{tz}(z')}$$

What are the approximations underlying the validity of the above solution ?

Problem P3.25

For an optically active medium, the wavenumbers for the right-hand and left-hand circularly polarized characteristic waves are k_r and k_l with $D_2/D_1 = \pm i$. In the xyz coordinate system, $\overline{D} = \hat{x} + \hat{y}i$ is right-hand circularly polarized when the wave is propagating in the $\hat{z}$ direction and left-hand polarized when the wave is in the $-\hat{z}$ direction. Consider a plane wave with

$$\overline{E} = \hat{y}E_0 e^{-ikz}$$

incident upon an anisotropic plasma layer with boundary surfaces at $z = 0$ and $z = -h$.

The solutions for the electric fields in all regions are assumed to be

$0 \le z$:

$$\overline{E} = \hat{y}E_0\left[e^{-ikz} + R_y e^{ikz}\right] + \hat{x}E_0 R_x e^{ikz}$$

$-h \le z \le 0$:

$$\overline{E} = (\hat{x}+\hat{y}i)E_0\left[Ae^{ik_r z}+Ce^{-ik_l z}\right]+(\hat{x}-\hat{y}i)E_0\left[Be^{-ik_r z}+De^{ik_l z}\right]$$

$z \le -h$:

$$\overline{E} = \hat{y}T_y E_0 e^{-ik_t z} + \hat{x}T_x E_0 e^{-ik_t z}$$

(a) Find the corresponding magnetic field components in all regions.

(b) Show that the wave amplitudes are obtained as follows (you can show by checking the continuity of the tangential field components at boundary surfaces):

$$R_x - iR_y = \frac{i}{D_r}\left\{(k_t + k_r)(k_l - k)e^{-ik_r h} + (k_t - k_l)(k_r + k)e^{ik_l h}\right\}$$

$$R_x + iR_y = \frac{i}{D_r}\left\{(k_t + k_l)(k_r - k)e^{-ik_l h} + (k_t - k_r)(k_l + k)e^{ik_r h}\right\}$$

$$D_r = (k_t + k_r)(k_l + k)e^{-ik_r h} + (k_t - k_l)(k_r - k)e^{ik_l h}$$

$$D_l = (k_t + k_l)(k_r + k)e^{-ik_l h} + (k_t - k_r)(k_l - k)e^{ik_r h}$$

$$A = \frac{i}{D_r}k(k_t - k_l)e^{ik_l h}$$

$$B = \frac{i}{D_l}k(k_t + k_l)e^{-ik_l h}$$

$$C = -\frac{i}{D_r}k(k_t + k_r)e^{-ik_r h}$$

$$D = -\frac{i}{D_l}k(k_t - k_r)e^{ik_r h}$$

$$T_x + iT_y = ik(k_r + k_l)\left\{\frac{1}{D_l}e^{i(k_r - k_l)h} - \frac{1}{D_r}e^{-i(k_r - k_l)h}\right\}$$

$$T_x - iT_y = k(k_r + k_l)\left\{\frac{1}{D_l}e^{i(k_r - k_l)h} + \frac{1}{D_r}e^{-i(k_r - k_l)h}\right\}$$

(c) Examine the limiting case as $k_t = k_r = k_l = k$; what are the values for all the wave amplitudes?

(d) Let $k_t = k$ and $D_r \simeq 4k^2 e^{-ik_r h}$ and $D_l \simeq 4k^2 e^{-ik_l h}$. Show that $R_x \pm iR_y = 0$ and

$$\frac{T_x}{T_y} = -\tan\left[\frac{(k_r - k_l)h}{2}\right]$$

(e) Let $k_t \to \infty$. Show that $T_x = T_y = 0$, $R_x \pm iR_y = 0$, and there is no rotation of the field vectors.

Problem P3.26

Show that for a coaxial line with outer radius b and inner radius a, the transmission line equations (3.4.16) can be derived in a similar manner. The inductance per unit length is $L = \mu \ln(b/a)/2\pi$ and the capacitance per unit length is $C = 2\pi\epsilon/\ln(b/a)$. The characteristic impedance is hence $A_0 = \eta \ln(b/a)/2\pi$. Find b/a for a 50 Ω coaxial line filled with a dielectric with $\epsilon = 8\epsilon_o$.

Problem P3.27

Show that for a transmission line terminated with load impedance Z_L at $z = 0$, the input impedance at $z = -l$ is

$$Z(-l) = Z_0 \frac{Z_{ln} - i\tan kl}{1 - iZ_{ln}\tan kl}$$

where $Z_{ln} = Z_L/Z_0$ and Z_0 is the characteristic impedance of the line. For an open circuit termination, $Z_{ln} \to \infty$. For a parallel-plate waveguide when $A_0 = (L/C)^{1/2}$, $L = \mu d/w$ and $C = \epsilon w/d$, show that in the quasi-static limit as $kl \ll 1$, the input impedance reduces to that of a capacitor, i.e., $Z(-l) = 1/j\omega C_0$ with $C_0 = \epsilon wl/d$.

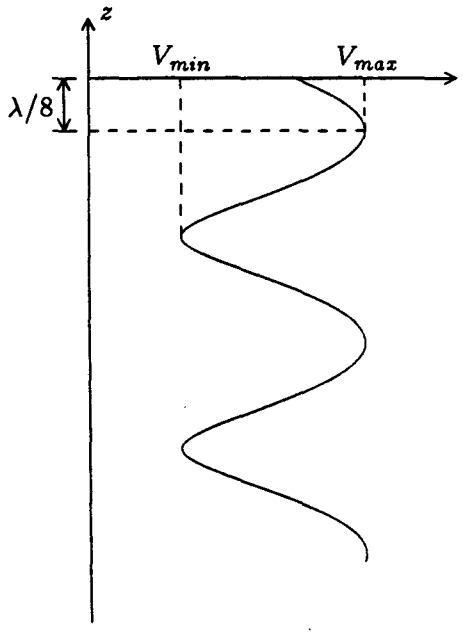


Figure P3.29

Problem P3.28

Plot the voltage standing wave pattern for a transmisison line with (i) open circiut termination $Z_{Ln} \to \infty$ (ii) short circuit termination $Z_{Ln} = 0$ and (iii) matched termination $Z_{Ln} = 1$.

Problem P3.29

A voltage standing wave pattern is shown in Figure P3.29. The voltage standing wave ratio is VSWR $= 3$. With characteristic impedance $Z_0 = 50\,\Omega$, show that the load impedance is $Z_L = 30 - i40\,\Omega$.

Problem P3.30

Consider the excitation of a parallel-plate waveguide by a current with $\overline{J}_s = \hat{x}J_s$, where J_s is a constant extending from $y = 0$ to $y = d$. Find the amplitudes of the excited modes when $d = 1$ cm and the free space wavelength is $\lambda = 1$ mm.

Problem P3.31

Consider the excitation of a parallel-plate waveguide by a current

sheet with

$$\overline{J}_s = \hat{x} J_s \cos \frac{3\pi x}{d}$$

Find the amplitudes of the excited modes.

Problem P3.32

A parallel-plate waveguide is excited by a line source placed at a distance h from the bottom plate. Write the fields produced by this source inside the waveguide as a superposition of the guided modes, and determine their mode amplitudes.

Problem P3.33

A plane slab of polystyrene ($\epsilon = 2.56\epsilon_o$) is 1 cm thick. At a frequency of 30 GHz, what are the propagating guided modes? What are their cutoff frequencies? Repeat the problem when the slab waveguide has air on one side and a dielectric medium with $\epsilon_t = 2\epsilon_o$ on the other.

Problem P3.34

Consider normal modes in a parallel-plate waveguide that is filled with an anisotropic medium. Let the medium in the waveguide be a cold plasma with a dc magnetic field $\overline{B}_0 = \hat{y} B_0$. Are the TE modes affected by the anisotropy of the plasma? Derive the TM mode functions.

Problem P3.35

In a parallel-plate waveguide, the region $z < 0$ is free space and the region $z > 0$ is filled with a dielectric medium having permittivity ϵ. How would you write the modes for both regions? For a TM mode incident upon the dielectric from the $z < 0$ region, show that there is no reflected wave when $f/f_c = \sqrt{(\epsilon + \epsilon_o)/\epsilon}$. Compare this result with the case of a wave incident upon a half-space medium at the Brewster angle.

Problem P3.36

The modes in a parallel-plate waveguide are orthogonal to each other because they are sinusoidal functions. In general we can prove that waveguide modes are orthogonal by showing that

$$\int_S dS \, (\overline{E}_m \times \overline{H}_n^*) \cdot \hat{z} = 0 \qquad \text{for} \quad m \neq n$$

where $\overline{E}_m$ is the $\overline{E}$ field for the mth mode, and $\overline{H}_n$ is the $\overline{H}$ field for the nth mode. The integral is carried out over the cross section of the waveguide. Show that the TE modes of a slab waveguide are orthogonal by verifying that

$$\int_{-\infty}^{\infty} dx\, E_{ym} H_{xn}^* = 0$$

Note that the integration range is from $-\infty$ to $+\infty$; field solutions both inside and outside the slab are needed in the calculations.

Problem P3.37

Although waves are usually guided by at least two plane interfaces, it is also possible to guide a wave with a single plane interface between two media. Such waves are called surface waves. Field components of a surface wave decay away from the interface exponentially. Consider a plane boundary surface at $z = 0$ separating two media with μ_0, ϵ_0 for $z > 0$ and μ_0, ϵ_p for $z < 0$.

(a) For TE waves with

$$\overline{E} = \hat{y} E_0 e^{ik_x x} \begin{cases} e^{-\alpha_0 z} & z > 0 \\ e^{\alpha_p z} & z < 0 \end{cases}$$

is it possible to have a TE surface wave? If so, find the dispersion relation.

(b) For TM waves with

$$\overline{H} = \hat{y} H_0 e^{ik_x x} \begin{cases} e^{-\alpha_0 z} & z > 0 \\ e^{\alpha_p z} & z < 0 \end{cases}$$

show that it is possible to have TM surface waves only if the permittivities of the upper and lower regions are real and opposite in sign.

(c) Assuming that

$$\epsilon_p = \epsilon_o \left[1 - \frac{\omega_p^2}{\omega^2}\right]$$

find the dispersion relation for the TM surface wave. (The dispersion relation should relate k_x to ω, and should contain no parameters other than ϵ_o, μ_o, and ω_p.)

(d) In the small ω limit, $\omega \to 0$, $\epsilon_p \to \infty$, and the plasma behaves like a perfect conductor. Verify that the dispersion relation reduces to the dispersion relation of a TM wave that propagates on the surface of a perfect conductor. What is α_0 in this case? Explain why such a wave is not called a surface wave.

(e) The incidence angle of an evanescent wave is complex,

$$\tan\theta = k_x/k_{0z} = k_x/i\alpha_0$$

Show that the angle of incidence for a plasma surface wave is equal to Brewster's angle.

Problem P3.38

Waves can be guided by continuously varying refractive indices as well as discrete layers.

(a) Show that the wave equation for electric field in an inhomogeneous, isotropic dielectric is

$$\nabla^2\overline{E} + \omega^2\mu\epsilon(\overline{r})\overline{E} + \nabla\left[\overline{E}\cdot\frac{\nabla\epsilon(\overline{r})}{\epsilon(\overline{r})}\right] = 0$$

(b) In a plane stratified medium, the permittivity $\epsilon(\overline{r}) = \epsilon(x)$ depends only on one coordinate. For a TE polarized wave of the form

$$\overline{E} = \hat{y}E_0 f(x)e^{ik_z z}$$

show that the wave equation in (a) reduces to

$$\left[\frac{d^2}{dx^2} + \omega^2\mu\epsilon(x)\right] f(x) = k_z^2 f(x)$$

Draw an analogy with the time-independent Schrödinger's equation for a particle subject to a potential $V(x)$

$$\left[\frac{d^2}{dx^2} - \frac{2m}{\hbar^2}V(x)\right]\psi(x) = -\frac{2m}{\hbar^2}E\,\psi(x)$$

where $\psi(x)$ is the probability wave function.

(c) Consider a medium in which $\epsilon = \epsilon_o[1 - (x/a)^2]$. Describe the corresponding quantum mechanical problem. Solve the wave equation by casting it into the following form:

$$\frac{d^2F(\xi)}{d\xi^2} + (\varsigma - \xi^2)F(\xi) = 0$$

and let $F(\xi) = H(\xi)e^{-\xi^2/2}$ to obtain

$$H''(\xi) - 2\xi H'(\xi) + (\varsigma - 1)H(\xi) = 0$$

Try a power series solution for $H(\xi)$

$$H(\xi) = \sum_{n=0} h_n \xi^n$$

Show that the series terminates at $n = N$ if $\varsigma = 2N + 1$. We then have a family of solutions $\{H_N(\xi); N = 0, 1, 2, \ldots\}$ corresponding to the eigenvalues $\varsigma = 1, 3, 5, \ldots$ $H_N(\varsigma)$ are called Hermite polynomials. Evaluate $H_N(\xi)$ for $N = 0, 1, 2$. The functions $H_N(\xi)e^{-\xi^2/2}$ are called Hermite Gaussians, which form a complete set of functions on $(-\infty, \infty)$. Find $f(x)$ and the cutoff frequencies of all TE modes.

(d) Sketch the dispersion diagram for TE_0 and TE_1 modes.

(e) Sketch $|\overline{E}|$ as a function of x for TE_0 and TE_1 modes. Note that the modes are localized around the origin. Therefore, these modes are approximately correct even if the permittivity does not go to negative values, as in the case of graded index fibers.

Problem P3.39

Consider a rectangular waveguide with dimensions $1\,\text{cm} \times 0.5\,\text{cm}$. What are the cutoff frequencies for the first five modes? If the waveguide is excited at 20 GHz, what are the propagation constants for the first five modes? If the waveguide is excited at 50 GHz, how many modes will propagate?

Problem P3.40

A rectangular waveguide is excited by a probe at the position $x = d$, $z = 0$. Assume that the probe extends from $y = 0$ to $y = b$, and approximate the current on the probe by

$$J(x, y, z) = I_0 \delta(x - d)\, \delta(z) \cos qy$$

What are the mode amplitudes that are excited by this probe? To achieve maximum excitation for the TE_{10} mode, where should the probe be placed?

Problem P3.41

The dispersion relation for an anisotropic medium in the kDB system is

$$\begin{bmatrix} u^2 - \nu\kappa & -i\nu\kappa_g \cos\theta \\ i\nu\kappa_g \cos\theta & u^2 - \nu(\kappa \cos^2\theta + \kappa_g \sin^2\theta) \end{bmatrix} \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} = 0$$

(a) For a wave propagating in the $\hat{z}$ direction coincidental with the dc magnetic field $\overline{B} = \hat{z}B$, find the polarizations and wavenumbers k_l and k_r for the two characteristic waves. Write the expressions for the electric field vector in the kDB system as well as in the xyz system.

(b) Repeat (a) for a wave propagating in the $-\hat{z}$ direction. Show that in the kDB system, the $\overline{D}$ vectors take the form

$$\overline{D}_l = D_1(\hat{e}_1 + \hat{e}_2 i)e^{ik_l z}$$
$$\overline{D}_r = D_1(\hat{e}_1 - \hat{e}_2 i)e^{ik_r z}$$

where $k_l = \omega/\sqrt{\nu(\kappa + \kappa_g)}$ and $k_r = \omega/\sqrt{\nu(\kappa - \kappa_g)}$ are in fact the wavenumber for, respectively, the right-hand and the left-hand circularly polarized characteristic waves propagating in the $-\hat{z}$ direction. In the xyz system, show that the $\overline{D}$ vectors take the same form as those in (a) with the exception of the negative sign in the exponent. Notice that here $\overline{k} = \hat{e}_3 k = -\hat{z}k$ and $\hat{e}_1 = \hat{x}$ and $\hat{e}_2 = -\hat{y}$.

(c) Consider a plane wave with

$$\overline{E} = \hat{y}E_0 e^{-ikz}$$

incident upon an anisotropic plasma layer with dc magnetic field in the $\hat{z}$ direction and the boundary surfaces at $z = 0$ and $z = -h$. The solutions for the electric fields in all regions are assumed to

be

$$0 \le z:$$
$$\overline{E} = \hat{y}E_0\left[e^{-ikz} + R_y e^{ikz}\right] + \hat{x}E_0 R_x e^{ikz}$$
$$-h \le z \le 0:$$
$$\overline{E} = (\hat{x}+i\hat{y})E_0\left[Ae^{ik_r z}+Ce^{-ik_r z}\right]+(\hat{x}-i\hat{y})E_0\left[Be^{-ik_l z}+De^{ik_l z}\right]$$
$$z \le -h:$$
$$\overline{E} = \hat{y}T_y E_0 e^{-ik_t z} + \hat{x}T_x E_0 e^{ik_t z}$$

Find the corresponding magnetic field components in all regions, and show that the wave amplitudes are obtained as follows:

$$R_x - iR_y = \frac{i}{D_r}\left\{(k_t - k_r)(k_r + k)e^{ik_r h}+(k_t + k_r)(k_r - k)e^{-ik_r h}\right\}$$
$$R_x + iR_y = \frac{-i}{D_l}\left\{(k_t - k_l)(k_l + k)e^{ik_l h} + (k_t + k_l)(k_l - k)e^{-ik_l h}\right\}$$
$$D_r = (k_t - k_r)(k_r - k)e^{ik_r h} + (k_t + k_r)(k_r - k)e^{-ik_r h}$$
$$D_l = (k_t - k_l)(k_l - k)e^{ik_l h} + (k_t + k_l)(k_l - k)e^{-ik_l h}$$
$$A = \frac{i}{D_r}k(k_t - k_r)e^{ik_r h}$$
$$C = \frac{-i}{D_r}k(k_t + k_r)e^{-ik_r h}$$
$$B = \frac{i}{D_l}k(k_t + k_l)e^{-ik_l h}$$
$$D = \frac{-i}{D_l}k(k_t - k_l)e^{ik_l h}$$
$$T_x + iT_y = i\frac{4kk_l}{D_l}$$
$$T_x - iT_y = -i\frac{4kk_r}{D_r}$$

(d) Examine the limiting case as $k_t = k_r = k_l = k$; what are the values for all the wave amplitudes?

(e) Let $k_t = k$ and $D_l \sim 4k^2 e^{-ik_l h}$ and $D_r \sim 4k^2 e^{-ik_r h}$. Show that

$$\frac{T_x}{T_y} = \tan\left[\frac{(k_r - k_l)h}{2}\right]$$

and

$$R_x \pm iR_y \approx 0$$

(f) Let $k_t \to \infty$, $D_l \sim D_r \to \infty$, and $T_x = T_y = 0$. Show that $R_x - iR_y = -ie^{i2k_r h}$, $R_x + iR_y \simeq ie^{i2k_l h}$, and

$$\frac{R_x}{R_y} = \tan[(k_r - k_l)h]$$

Problem P3.42

Consider that the rectangular waveguide is partially filled with a dielectric medium ϵ_1 from $y = 0$ to $y = d$, and a dielectric medium ϵ_2 from $y = d$ to $y = b$. Show that the guided waves inside are hybrid modes. Letting $\epsilon_1 \approx \epsilon_2$, show that the cutoff frequency for the dominant mode is

$$\omega_c \approx \frac{\pi}{b}\left[\frac{\epsilon_1(b-d) + \epsilon_2 d}{\mu\epsilon_1\epsilon_2 b}\right]^{1/2}$$

Problem P3.43

In a fiberglass waveguide having a center core with a radius of the order of 1 μm and cladding with a radius of the order of 100 μm, the HE_{11} mode operating range can be extended to the visible range if the refractive indices $n = c\sqrt{\mu\epsilon}$ and $n_1 = c\sqrt{\mu_1\epsilon_1}$ are also very close. Because the cladding is very thick in comparison to the core, the wave guidance by an optical fiber can be treated with the dielectric waveguide model. Find the value of $(n^2 - n_1^2)^{1/2}$, called the *numerical aperture*, so that the cutoff frequency for the next higher-order mode will be 6×10^{14} Hz. When fiberglass is used as the transmission medium in communications, it provides not only large bandwidth and channel capacity but also physical compactness and flexibility. Compare the result with the slab and with the metallic waveguides.

Problem P3.44

Consider a coaxial line with inner radius a and outer radius b. Assume that $b = a(1 + \delta)$ and $\delta \ll 1$. The fundamental mode in this waveguide is TEM, which has zero cutoff wavenumber. What are the cutoff wavenumbers for the higher-order modes? Rigorously this

requires the solution of the boundary value problem in cylindrical coordinates. The exact solution will involve Bessel and Neumann functions. However, the question can be answered by observing that, when δ is very small, the guiding space can be modeled as that between two parallel plates with $x = \rho\phi$ and y from a to b. Notice that $x = 2\pi a \approx 2\pi b$ and that the fields must be the same at $x = 0$. Thus there is a periodic variation in the $\hat{x}$ direction. Using this model, show that the cutoff wavenumber of the next higher order mode is $k_c \approx 1/a$. Note the similarity to the cutoff wavenumber for the TE_{20} mode in a rectangular waveguide with width $2\pi a$ and height δa. Confirm the answer by evaluating the cutoff wavenumber from the guidance condition. Note that the cutoff wavenumber for TE_1 and TM_1 modes in a parallel-plate waveguide is $k_c = \pi/\delta a$, which is much larger than $1/a$.

Problem P3.45

Consider a parallel plate waveguide filled with plasma medium for $z > 0$ and with dielectric medium for $z < 0$. The plasma medium has permittivity

$$\epsilon_p = \epsilon_o \left[1 - \frac{\omega_p^2}{\omega^2}\right]$$

and the dielectric medium has permittivity $\epsilon_1 = 3\epsilon_o$. Let

$$d = \sqrt{5/3}\,\frac{3\pi}{\omega\sqrt{\mu_o\epsilon_o}}$$

(a) Consider the case when $\epsilon_p = \epsilon_1$ (namely the whole waveguide is filled with dielectric). How many propagating TM modes can be guided?

(b) Let $\omega_p = (1/2)\omega$. For waves propagating in the $+z$ direction, which of the above TM modes will be totally reflected at the dielectric-plasma boundary? Why?

(c) One of the above TM modes will be totally transmitted (no reflection). Which one and why?

(d) For a given excitation frequency ω, at what plasma frequency ω_p will all of the above TM modes be totally reflected?

Problem P3.46

Consider an optical fiber with radius a and permittivity ϵ imbedded in a cladding which can be regarded as having infinite radius and with permittivity ϵ_1. Assume $\delta = 1 - \epsilon_1/\epsilon \ll 1$.

(a) Write all field components of $\overline{E}$ and $\overline{H}$ for $\rho \leq a$ and $\rho \geq a$ for general $\epsilon_1 < \epsilon$. Under the assumption of $\delta \ll 1$, show that with $E_y = E_\rho \sin\phi + E_\phi \cos\phi$ and $H_x = H_\rho \cos\phi - H_\phi \sin\phi$ can be approximated by, for $\rho \leq a$

$$E_y = -E\frac{k_z k}{\omega\epsilon}\left[J_1'(k_\rho\rho)\sin^2\phi + \frac{1}{k\rho}J_1(k_\rho\rho)\cos^2\phi\right]e^{ik_z z}$$
$$H_x = -\frac{\omega\epsilon}{k_z}E_y$$

and for $\rho \geq a$

$$E_{1y} = -E_1\frac{k_z k_1}{\omega\epsilon_1}\left[H_1^{(1)\prime}(k_\rho\rho)\sin^2\phi + \frac{1}{k_{1\rho}\rho}H_1^{(1)}(k_{1\rho}\rho)\cos^2\phi\right]$$
$$H_{1x} = -\frac{\omega\epsilon_1}{k_z}E_{1y}$$

(b) Show that, by making use of the recurrence relation for Bessel functions, the terms multiplied by $\cos 2\phi$ are negligible, i.e.

$$E_y = -E\frac{k_z k}{\omega\epsilon}J_0(k_\rho\rho)e^{ik_z z}$$
$$E_{1y} = -E_1\frac{k_z k_1}{\omega\epsilon_1}H_0(k_\rho\rho)e^{ik_z z}$$

(c) Show that the following approximation is numerically accurate.

$$\left.\begin{array}{ll} \rho \leq a & J_0(k_\rho\rho) \\ a < \rho & \dfrac{J_0(k_\rho a)}{H_0^{(1)}(k_{1\rho}a)}H_0^{(1)}(k_{1\rho}\rho) \end{array}\right\} \approx \exp\left[-\frac{1}{2}\left(\frac{\rho}{\rho_0}\right)^2\right]$$

where

$$\rho_0^2 = a^2/\ln\left[(k^2 - k_1^2)a^2\right]$$

is called the spot size.

(d) Show that Poynting's power density in the $\hat{z}$ direction is

$$S_z(\rho) = |E|^2 \frac{k_z k_{1\rho}^2}{8\omega\epsilon} e^{-\rho^2/\rho_0^2}$$

and the ratio of the power carried by the core to the total power transfered along the fiber is

$$P_{core}/P_{total} = 1 - e^{-a^2/\rho_0^2}$$

(e) Let the spot size equal the core radius, $\rho_0 = a$. Find $(k^2 - k_1^2)^{1/2}a$ and show that it is below the cutoff of the next higher mode. Calculate the spot size for $(k^2 - k_1^2)^{1/2}a = 0.8$.

Problem P3.47

Consider a rectangular cavity with a square base and height a. A thin slab of dielectric with thickness d is placed on the base. Determine the change in resonant frequency by using the perturbation formula. Show that, if we assume unperturbed values for the fields, the result is

$$\frac{\omega - \omega_0}{\omega_0} \approx -\frac{1}{2}(\epsilon_r - 1)\frac{d}{a}$$

Problem P3.48

Derive a material perturbation formula for propagation constants in a waveguide. Show that the following formula is exact:

$$k_z - k_{oz} = \omega \frac{\iint (\Delta\epsilon \overline{E} \cdot \overline{E}_0^* + \Delta\mu \overline{H} \cdot \overline{H}_0^*) ds}{\iint_s (\overline{E}_0^* \times \overline{H} + \overline{E} \times \overline{H}_0^*) \cdot \hat{z}\, ds}$$

where k_{oz} denotes the unperturbed propagation constant. The field vectors in this formula have no dependence on z. The fields can be approximated by their unperturbed values outside the material and the quasi-static values inside the material.

Consider a circular waveguide of radius a containing a concentric rod of radius b. Show that the change in propagation constant is given by

$$k_z - k_{oz} = k \frac{2.146}{\sqrt{1 - (k_c/k)^2}} \frac{\epsilon_r - 1}{\epsilon_r + 1} \left(\frac{b}{a}\right)^2$$

Problem P3.49

The field in a spherical cavity with radius a is given by

$$H_\phi = \frac{H_0}{kr} J_1\left(2.744\frac{r}{a}\right)\sin\theta$$

Assume that the cavity is perturbed by a concentric dielectric with radius a and permittivity ϵ. Prove that

$$\frac{\omega - \omega_0}{\omega_0} \approx -0.291\frac{\epsilon_r - 1}{\epsilon_r + 2}\left(2.744\frac{b}{a}\right)^3$$

by using the quasi-static field inside the dielectric sphere; this is approximately $[3/(2+\epsilon_r)]E_0$, where $\epsilon_r = \epsilon/\epsilon_o$ and E_0 is the unperturbed field E_0 outside the dielectric.

IV

RADIATION

4.1 Čerenkov Radiation
4.2 Green's Functions
4.3 Hertzian Dipoles
a. Hertzian Electric Dipole
b. Hertzian Magnetic Dipole and Small Loop Antenna
4.4 Radiation Fields
a. Radiation Resistance
4.5 Biconical Antennas
a. Formulation and Wave Solutions
b. Solution in the Air Region and Dipole Fields
c. Solution in the Antenna Region
d. Transmission Line Model
e. Formal Solution of Biconical Antenna Problem
4.6 Linear Antenna Arrays
a. Uniform Array Antenna with Progressive Phase Shift
b. Array Antenna with Nonuniform Current Distributions
c. Dolph-Chebyshev Arrays
d. Array Pattern Synthesis
4.7 Contour Integration Methods
a. Cauchy's Theorem
b. Asymptotic Series for Hankel Functions
c. Saddle-Point Method
4.8 Integral Formulations for Dipoles in Layered Media
4.9 Dipole on One-Layer Medium
4.10 Dipole Above Two-Layer Medium
Problems

4.1 Čerenkov Radiation

In 1934, Čerenkov discovered experimentally that all liquids and solids emit visible radiation when bombarded by fast-moving electron beams. He discovered that (i) in order to achieve radiation, the velocity of the electrons must be very large, (ii) the angles of radiation are related to the velocity of the beam, and (iii) the emitted light has the electric field vector polarized parallel to the plane determined by the direction of the beam and the direction of the radiation. Many unsuccessful attempts were made to explain the discovery with various microscopic approaches. In 1937, Frank and Tamm used the macroscopic theory and established that an electron moving uniformly in a medium characterized by a refractive index larger than unity radiates light if the electron velocity is greater than the velocity of light in the medium. Because the discovery of this phenomenon, known as *Čerenkov radiation*, marked a significant triumph for macroscopic electromagnetic theory, we shall devote this section to this subject.

The source of radiation is a particle with charge q moving at a velocity $\overline{v}$ in an isotropic medium. The velocity will decrease as a result of radiation. To simplify discussions, we assume that $\overline{v}$ is a constant in the direction $\hat{z}$. The current density of the moving charge is

$$\overline{J}(\overline{r},t) = \hat{z}qv\delta(x)\delta(y)\delta(z-vt)$$

In the cylindrical coordinate system, we have ϕ symmetry. Noticing that

$$\int d\rho\,\delta(\rho) = 1 = \iint dx\,dy\,\delta(x)\delta(y) = \int 2\pi\rho d\rho\,\delta(x)\delta(y)$$

We can write $\delta(x)\delta(y) = \delta(\rho)/2\pi\rho$ and therefore

$$\overline{J}(\overline{r},t) = \hat{z}qv\delta(z-vt)\frac{\delta(\rho)}{2\pi\rho} \tag{1}$$

This source is not time harmonic. We transform to the frequency domain and obtain

$$\overline{J}(\overline{r},\omega) = \frac{1}{2\pi}\int dt\,\overline{J}(\overline{r},t)e^{i\omega t} = \hat{z}\frac{q}{4\pi^2\rho}e^{i\omega z/v}\,\delta(\rho) \tag{2}$$

For each spectrum component ω we solve for the electric field

$$\overline{E}(\overline{r}) = \frac{1}{2\pi}\int dt\,\overline{E}(\overline{r},t)e^{i\omega t}$$

The time-domain values are obtained by the inverse Fourier transform:

$$\overline{E}(\bar{r},t) = \int d\omega\, \overline{E}(\bar{r}) e^{-i\omega t} \tag{3}$$

The governing equation for the electric field becomes

$$\nabla \times \nabla \times \overline{E}(\bar{r}) - k^2 \overline{E}(\bar{r}) = \hat{z} \frac{i\omega\mu q}{4\pi^2 \rho} e^{i\omega z/v}\, \delta(\rho) \tag{4}$$

This equation is conveniently solved by defining a vector Green's function $\overline{g}(\rho, z)$ such that

$$\overline{E}(\bar{r}) = \left[\overline{\overline{I}} + \frac{1}{k^2}\nabla\nabla\right] \cdot \overline{g}(\rho, z) = \overline{g}(\rho, z) + \frac{1}{k^2}\nabla[\nabla \cdot \overline{g}(\rho, z)] \tag{5}$$

We obtain from (4) the wave equation for $\overline{g}(\rho, z)$:

$$(\nabla^2 + k^2)\overline{g}(\rho, z) = -\hat{z} \frac{i\omega\mu q}{4\pi^2 \rho} e^{i\omega z/v}\, \delta(\rho) \tag{6}$$

In view of the z dependence on the right-hand side and the azimuthal symmetry of the problem, we write the wave equation in the cylindrical coordinate system. Let

$$\overline{g}(\rho, z) = \hat{z} g(\rho) \frac{i\omega\mu q}{2\pi} e^{i\omega z/v} \tag{7}$$

Then we obtain

$$\left[\frac{1}{\rho}\frac{d}{d\rho}\left(\rho \frac{d}{d\rho}\right) - \frac{\omega^2}{v^2} + k^2\right] g(\rho) = -\frac{\delta(\rho)}{2\pi\rho} \tag{8}$$

For $\rho \neq 0$, the above equation becomes

$$\left[\frac{1}{\rho}\frac{d}{d\rho}\left(\rho \frac{d}{d\rho}\right) + k_\rho^2\right] g(\rho) = 0$$

where

$$k_\rho = \sqrt{k^2 - \frac{\omega^2}{v^2}}$$

This is the Bessel equation of zeroth order. Since (8) exhibits a singularity at $\rho = 0$ and the solution to the Bessel equation should represent an outgoing wave, we choose

$$g(\rho) = CH_0^{(1)}(k_\rho \rho) \tag{9}$$

The constant C is determined by matching the boundary condition at $\rho \to 0$. Integrating (8) over an infinitesimal area $2\pi\rho d\rho$ and letting $\rho \to 0$, we have

$$\lim_{\rho \to 0} 2\pi\rho \frac{dg(\rho)}{d\rho} = -1$$

Using the asymptotic formula for $H_0^{(1)}(k_\rho\rho) \approx i(2/\pi)\ln(k_\rho\rho)$, we obtain $C = i/4$ and from (9)

$$g(\rho) = \frac{i}{4} H_0^{(1)}(k_\rho \rho) \tag{10}$$

This is the scalar Green's function in cylindrical coordinates. For two-dimensional problems independent of z, the scalar Green's function is simply (10) with $k_\rho = k$.

The solution for the electric field is determined from (7) and (5).

$$\overline{E}(\bar{r}) = \frac{-q}{8\pi\omega\epsilon} \left[\hat{z}k^2 + i\frac{\omega}{v}\nabla\right] H_0^{(1)}(k_\rho\rho) e^{i\omega z/v} \tag{11}$$

Since we are interested in radiation from the charge, we use the asymptotic values of $H_0^{(1)}(k_\rho\rho)$ to find the far-field solutions. In the radiation zone, $k_\rho\rho \gg 1$, and $H_0^{(1)}(k_\rho\rho) \approx \sqrt{2/i\pi k_\rho\rho}\, e^{ik_\rho\rho}$. We get

$$\overline{E}(\bar{r}) \approx \frac{q}{8\pi\omega\epsilon}\sqrt{\frac{2k_\rho}{i\pi\rho}} \left[\hat{\rho}\frac{\omega}{v} - \hat{z}k_\rho\right] e^{i(k_\rho\rho + \omega z/v)} \tag{12}$$

This represents a plane wave with wave vector $\bar{k} = \hat{\rho}k_\rho + \hat{z}\omega/v$, provided that k_ρ is real.

All the phenomena observed by Čerenkov can be explained with (12).

(i) We see that k_ρ is real if $\omega^2\mu\epsilon > \omega^2/v^2$, i.e.,

$$v > \frac{1}{\sqrt{\mu\epsilon}} = \frac{c}{n} \tag{13}$$

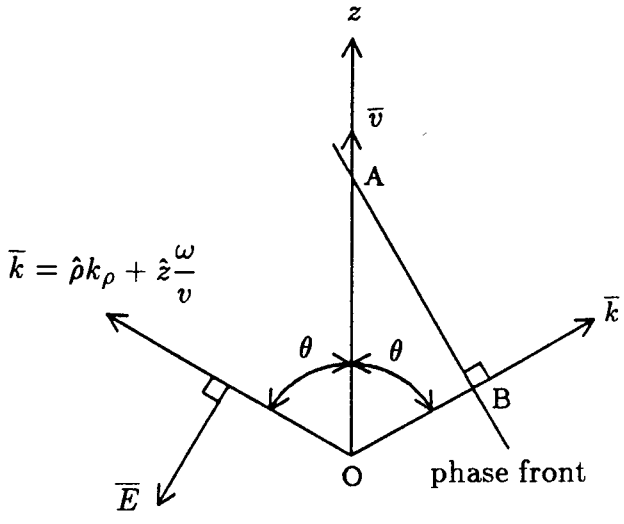


Figure 4.1.1 Čerenkov radiation.

Thus, plane waves are radiated if the velocity of the charge is larger than the velocity of light in the medium. When the charge velocity v is smaller than the light velocity, k_ρ is imaginary and the wave is evanescent in the $\hat{\rho}$ direction.

(ii) The constant phase front of the plane waves forms a cone around the $\hat{z}$ direction. The direction θ that $\overline{k}$ makes with $\hat{z}$ [Fig. 4.1.1] is determined from

$$\cos\theta = \frac{\omega}{kv} = \frac{1}{n\beta} \tag{14}$$

where $\beta = v/c$. Note that θ has a real value only if $n\beta > 1$. Notice from Figure 4.1.1 that for the phase front represented by AB, OA is the distance traveled by the charge particle and OB is the distance traveled by the wave originated from point O.

(iii) With regard to the polarization of the emitted electromagnetic wave, we observe from (12) that $\overline{E}$ lies in the plane determined by $\overline{k}$ and $\hat{z}$ [Fig. 4.1.1]. It is clear the $\overline{E}$ is also perpendicular to the $\overline{k}$ vector because $\overline{k}\cdot\overline{E} = 0$. We can attempt an explanation by imagining that when a charged particle is at rest the electric fields point radially. When it is moving, the field lines are bent at an angle proportional to

the speed of the particle. As the speed exceeds the velocity of light in the medium, the field lines can actually break away from the charge and constitute a radiated wave. Therefore, the electric field vector lies in the plane determined by the direction of motion and the direction of radiation.

To calculate radiated power, we first compute the magnetic field from Faraday's law, which gives

$$\overline{H} = \hat{\phi}\frac{q}{8\pi}\sqrt{\frac{2k_\rho}{i\pi\rho}}\, e^{i(k_\rho\rho+\omega z/v)} \tag{15}$$

when terms of order $\rho^{-3/2}$ are neglected.

Consider a cylinder of length l and radius ρ. The total energy radiated through the surface of the cylinder is given by

$$S_\rho = 2\pi\rho l\int_{-\infty}^{\infty} dt\,[\overline{E}(\overline{r},t)\times\overline{H}(\overline{r},t)]_\rho \tag{16}$$

In order to calculate S_ρ, we first determine the real, space–time $\overline{E}$ and $\overline{H}$ fields. Since only the ρ component of Poynting's vector is required, we need the field component E_z. We use the inverse Fourier transform (3). Noting that k_ρ is proportional to ω because $k_\rho{}^2 = \omega^2\mu\epsilon - \omega^2/v^2$, we change the integration limits from $(-\infty,\infty)$ to $(0,\infty)$ and obtain

$$E_z = -\frac{q}{4\pi}\sqrt{\frac{2}{\pi\rho}}\int_0^\infty d\omega\, k_\rho\frac{\sqrt{k_\rho}}{\omega\epsilon}\cos(\omega t - k_\rho\rho - \frac{\omega z}{v} + \frac{\pi}{4}) \tag{17}$$

By the same token,

$$H_\phi = \frac{q}{4\pi}\sqrt{\frac{2}{\pi\rho}}\int_0^\infty d\omega\,\sqrt{k_\rho}\cos(\omega t - k_\rho\rho - \frac{\omega z}{v} + \frac{\pi}{4}) \tag{18}$$

Now we are able to calculate S_ρ, which is a triple integral

$$S_\rho = \frac{q^2 l}{4\pi^2}\int_{-\infty}^{\infty} dt\int_0^\infty d\omega\int_0^\infty d\omega'\, k_\rho\frac{\sqrt{k_\rho}}{\omega\epsilon}\sqrt{k'_\rho}$$
$$\cos(\omega t - k_\rho\rho - \frac{\omega z}{v} + \frac{\pi}{4})\cos(\omega' t - k'_\rho\rho - \frac{\omega' z}{v} + \frac{\pi}{4})$$

where $k'_\rho = \omega'(\mu\epsilon - 1/v^2)^{1/2}$. First we integrate with respect to t. Let $\alpha = k_\rho \rho/\omega + z/v$, we observe that

$$\int_{-\infty}^{\infty} dt \cos\left[\omega'(t+\alpha) + \frac{\pi}{4}\right] \cos\left[\omega(t+\alpha) + \frac{\pi}{4}\right]$$
$$= \frac{1}{2}\int_{-\infty}^{\infty} dt \cos[(\omega - \omega')(t+\alpha)] = \pi\delta(\omega - \omega') \quad (19)$$

where we make use of the delta function

$$\delta(\omega - \omega') = \frac{1}{2\pi}\int_{-\infty}^{\infty} dt\, e^{i(\omega-\omega')t}$$

Thus we obtain

$$S_\rho = \frac{q^2 l}{4\pi}\int_0^\infty d\omega \frac{k_\rho^2}{\omega\epsilon} = \frac{\mu q^2 l}{4\pi}\int_0^\infty d\omega\, \omega\left[1 - \frac{1}{n^2\beta^2}\right] \quad (20)$$

Even though the integration limit is from 0 to ∞, we must remember that the above result is valid only for $n^2 > 1/\beta^2$ in order to achieve Čerenkov radiation. Since all materials are dispersive, the integration limit is actually determined by the frequency range of the refractive index n for which the Čerenkov radiation condition is satisfied. With the use of the above equation, energy radiated per unit length of the electron path can be calculated for materials of various refractive indices. Furthermore, we must note that the above theoretical treatment assumes a constant velocity v. As the charge radiates, the particle slows down and eventually ceases to radiate as $\beta^2 \leq 1/n^2$.

4.2 Green's Functions

In antenna and radiation problems, we are interested in finding the solution of electromagnetic fields in the presence of a source distribution $\overline{J}(\overline{r})$ and $\rho(\overline{r})$. The current and charge distributions are related by the conservation law. For time-harmonic fields, $i\omega\rho(\overline{r}) = \nabla \cdot \overline{J}(\overline{r})$. From Maxwell's equations

$$\nabla \times \overline{E}(\overline{r}) = i\omega\mu\overline{H}(\overline{r}) \quad (1)$$
$$\nabla \times \overline{H}(\overline{r}) = -i\omega\epsilon\overline{E}(\overline{r}) + \overline{J}(\overline{r}) \quad (2)$$

we can eliminate $\overline{H}(\bar{r})$ and obtain the following equation for $\overline{E}(\bar{r})$:

$$\nabla \times \nabla \times \overline{E}(\bar{r}) - k^2 \overline{E}(\bar{r}) = i\omega\mu \overline{J}(\bar{r}) \tag{3}$$

where $k^2 = \omega^2 \mu\epsilon$. To determine $\overline{E}$ in terms of the given source $\overline{J}$, we introduce the use of dyadic Green's functions. A more traditional approach can be taken by making use of a vector potential which is described in Problem P4.1.

A Green's function is the response due to a point source and is useful in expressing a field in terms of its source. Since $\overline{E}(\bar{r})$ is a vector and so is $\overline{J}(\bar{r})$, we write

$$\overline{E}(\bar{r}) = i\omega\mu \iiint dV' \, \overline{\overline{G}}(\bar{r}, \bar{r}') \cdot \overline{J}(\bar{r}') \tag{4}$$

where $\overline{\overline{G}}(\bar{r}, \bar{r}')$ is the dyadic Green's function that enables one to determine the electric field $\overline{E}$ from a given distribution $\overline{J}$.

Since the electric field $\overline{E}$ is a vector and the current source $\overline{J}$ is also a vector, the Green's function $\overline{\overline{G}}$ must be a dyad that operates on a vector giving rise to another vector. The operation may be thought of as a square matrix representing $\overline{\overline{G}}$ multiplying a column matrix representing $\overline{J}$, giving rise to another column matrix representing $\overline{E}$. From another point of view, we can define a dyad from two vectors. Let $\overline{A}$ and $\overline{B}$ be two vectors. The dot product $\overline{A} \cdot \overline{B}$ gives a scalar and the cross product $\overline{A} \times \overline{B}$ gives a vector. Introducing a third vector $\overline{C}$ and making use of the identity $\overline{B} \times (\overline{A} \times \overline{C}) = \overline{A}\,\overline{B} \cdot \overline{C} - \overline{C}\,\overline{B} \cdot \overline{A}$, we can identify the direct product $\overline{A}\,\overline{B}$ as a dyad $\overline{\overline{D}} = \overline{A}\,\overline{B}$. In index notation, the ijth component of $\overline{\overline{D}}$ is $D_{ij} = A_i B_j$. The operation $\overline{\overline{D}} \cdot \overline{C}$ yields the vector $\overline{A}(\overline{B} \cdot \overline{C})$, which is the vector $\overline{A}$ weighted by the scalar $\overline{B} \cdot \overline{C}$. In index notation the ith component of $\overline{\overline{D}} \cdot \overline{C}$ is $D_{ij} C_j = A_i B_j C_j$ where repeated index j implies summation from 1 to 3. Another example is obtained by considering $\nabla \times \nabla \times \overline{E} = \nabla\nabla \cdot \overline{E} - \nabla^2 \overline{E}$. The operator $\nabla\nabla$ is now a dyadic operator.

The right-hand side of (3) can be cast in a form similar to (4) by using the three-dimensional delta function $\delta(\bar{r} - \bar{r}')$ such that

$$\overline{J}(\bar{r}) = \iiint dV' \delta(\bar{r} - \bar{r}') \overline{\overline{I}} \cdot \overline{J}(\bar{r}') \tag{5}$$

where $\overline{\overline{I}}$ is a unit dyad which can be represented by a unit diagonal matrix. The operation of $\overline{\overline{I}}$ on any vector yields the vector itself.

Substituting (4) and (5) into (3) and noting that the integral holds for arbitrary $\bar{J}(\bar{r}')$, we obtain a differential equation for the dyadic Green's function $\bar{\bar{G}}(\bar{r},\bar{r}')$

$$\nabla \times \nabla \times \bar{\bar{G}}(\bar{r},\bar{r}') - k^2\bar{\bar{G}}(\bar{r},\bar{r}') = \bar{\bar{I}}\delta(\bar{r}-\bar{r}') \tag{6}$$

The interchange of the differential operator $\nabla\times\nabla\times$ and the volume integral in (4) has serious implications when $\bar{r}$ is inside the source region, which will be discussed in Chapter VI. In this chapter the observation point $\bar{r}$ is always assumed to be outside the source distribution.

The dyadic Green's function can in turn be expressed in terms of a scalar Green's function $g(\bar{r},\bar{r}')$

$$\bar{\bar{G}}(\bar{r},\bar{r}') = \left[\bar{\bar{I}} + \frac{1}{k^2}\nabla\nabla\right] g(\bar{r},\bar{r}') \tag{7}$$

Here we make use of the dyadic operator $\nabla\nabla$. We introduce (7) in (6) and notice that the operator $\nabla\times\nabla\times$ yields zero when operated on the second term in (7) because the curl of a gradient is zero, $\nabla\times\nabla\times\nabla\nabla = \nabla\times(\nabla\times\nabla)\nabla = 0$. Also $\nabla\times\nabla\times(\bar{\bar{I}}g) = \nabla\nabla g - \bar{\bar{I}}\nabla^2 g$. We obtain the following differential equation for $g(\bar{r},\bar{r}')$:

$$(\nabla^2 + k^2)g(\bar{r},\bar{r}') = -\delta(\bar{r}-\bar{r}') \tag{8}$$

It is seen that Green's functions are responses to point sources. We now determine the scalar Green's function $g(\bar{r},\bar{r}')$ from (8) in the spherical coordinate system.

We first translate the coordinate origin such that $\bar{r}' = 0$. We write

$$\nabla\cdot\nabla g(\bar{r}) + k^2 g(\bar{r}) = -\delta(\bar{r}) \tag{9}$$

We see that (9) and its solution for $g(\bar{r})$ are spherically symmetric and independent of θ and ϕ. In the spherical coordinate system, (9) becomes

$$\frac{1}{r}\frac{d^2}{dr^2}[rg(\bar{r})] + k^2 g(\bar{r}) = -\delta(\bar{r})$$

For $r \neq 0$ the right-hand side is zero and we have

$$\frac{d^2}{dr^2}[rg(\bar{r})] + k^2 rg(\bar{r}) = 0 \tag{10}$$

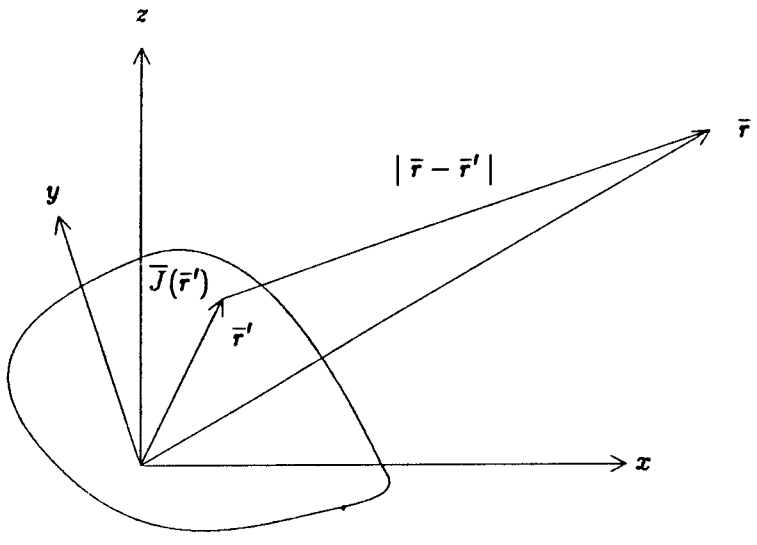


Figure 4.2.1 Observation point $\bar{r}$ is outside the source region.

The solution must represent an *outgoing* wave. We find

$$g(r) = C\frac{e^{ikr}}{r} \tag{11}$$

The constant C is determined by integrating (9) over a sphere of infinitesimal radius δ centered at the origin. In view of Gauss' theorem in vector calculus, integration of the first term of (9) yields

$$\iiint dV\, \nabla^2 g = \oiint_{r=\delta} dS\, \hat{r} \cdot \nabla g = \left[4\pi r^2 \frac{dg(\bar{r})}{dr}\right]_{r=\delta} \tag{12}$$

We obtain from (9)

$$\left[4\pi r^2 \frac{dg(\bar{r})}{dr}\right]_{r=\delta} + k^2 \int_0^{\delta} dr\, 4\pi r^2 g(\bar{r}) = -1 \tag{13}$$

Introducing (11) we see that in the limit of $\delta \to 0$, the second term is proportional to δ^2 and vanishes. The first term gives $-4\pi C$. Thus we find $C = 1/4\pi$.

Notice that r is the distance between the source and the observation point. Transforming back to the original coordinate system, it is seen that the distance r becomes $|\bar{r}-\bar{r}'|$. We obtain the scalar Green's function

$$g(\bar{r},\bar{r}') = \frac{e^{ik|\bar{r}-\bar{r}'|}}{4\pi\,|\bar{r}-\bar{r}'|} \tag{14}$$

where $|\bar{r}-\bar{r}'|$ is the distance between the field point $\bar{r}$ and the source point $\bar{r}'$ [Fig. 4.2.1]. We assume the observation point is outside of the source region. Substituting (7) into (4) and noting that the integration is over the primed quantities while the del operators operate only on the unprimed quantities, we can take the operators out of the integral and write

$$\bar{E}(\bar{r}) = i\omega\mu\left[\bar{\bar{I}} + \frac{1}{k^2}\nabla\nabla\right]\cdot\iiint dV'\,g(\bar{r},\bar{r}')\bar{J}(\bar{r}') \tag{15}$$

In terms of the scalar Green's function in spherical coordinates as determined in (14), we find

$$\bar{E}(\bar{r}) = i\omega\mu\left[\bar{\bar{I}} + \frac{1}{k^2}\nabla\nabla\right]\cdot\iiint dV'\frac{e^{ik|\bar{r}-\bar{r}'|}}{4\pi\,|\bar{r}-\bar{r}'|}\bar{J}(\bar{r}') \tag{16}$$

Thus, for a prescribed source distribution $\bar{J}(\bar{r}')$ in an unbounded isotropic medium, the electric field is determined by evaluating the integral (16). The magnetic field is then calculated from Faraday's law in (1).

4.3 Hertzian Dipoles

a. Hertzian Electric Dipole

The most fundamental model for radiating structures is a Hertzian electric dipole which consists of a current-carrying element with an infinitesimal length l. Denoting the current dipole moment with Il, the current density $\bar{J}(\bar{r})$ of a Hertzian dipole pointing in the $\hat{z}$ direction and located at the origin [Fig. 4.3.1] is

$$\bar{J}(\bar{r}') = \hat{z}Il\,\delta(\bar{r}') \tag{1}$$

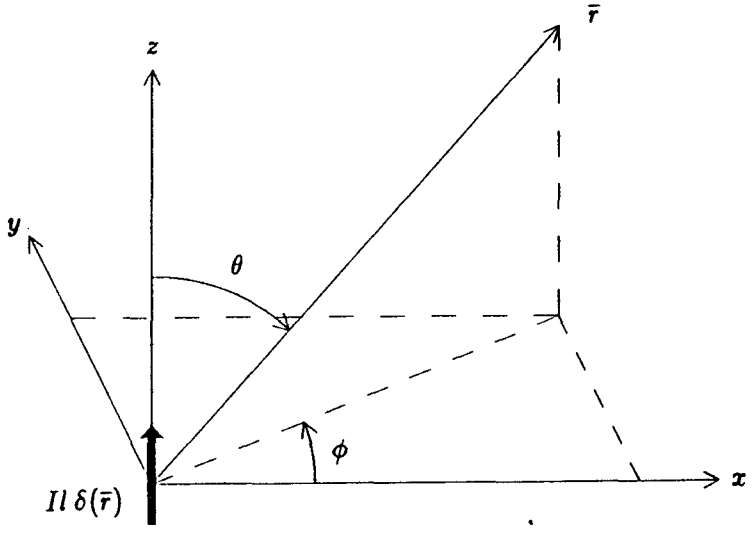


Figure 4.3.1 Hertzian electric dipole.

A Hertzian dipole can be modeled as two charge reservoirs of equal and opposite charge q and separated by an infinitesimal distance l. One may think of two conducting spheres or a capacitor connected by a constant current source. The dipole has moment $p = ql$ and oscillates in time with angular frequency ω. The current dipole moment is thus $Il = -i\omega p$.

If we notice that with the scalar Green's function $g(r) = e^{ikr}/4\pi r$ and $\partial g(r)/\partial z = (ik - 1/r)\cos\theta\, g(r)$, the electric field due to the Hertzian dipole is calculated to be

$$\begin{aligned}
\overline{E}(\bar{r}) &= i\omega\mu\left[\overline{\overline{I}} + \frac{1}{k^2}\nabla\nabla\right]\cdot\iiint dV'\frac{e^{ik|\bar{r}-\bar{r}'|}}{4\pi|\bar{r}-\bar{r}'|}\hat{z}Il\delta(\bar{r}') \\
&= i\omega\mu Il\left[\hat{z} + \frac{1}{k^2}\nabla\frac{\partial}{\partial z}\right]\frac{e^{ikr}}{4\pi r} \\
&= i\omega\mu Il\Bigg\{(\hat{r}\cos\theta - \hat{\theta}\sin\theta) \\
&\quad + \frac{1}{k^2}\left[\hat{r}\frac{\partial}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial}{\partial\theta} + \hat{\phi}\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\right]\left[ik - \frac{1}{r}\right]\cos\theta\Bigg\}\frac{e^{ikr}}{4\pi r}
\end{aligned}$$

$$= -i\omega\mu Il\frac{e^{ikr}}{4\pi r}\left\{\hat{r}\left[\frac{i}{kr}+\left(\frac{i}{kr}\right)^2\right]2\cos\theta + \hat{\theta}\left[1+\frac{i}{kr}+\left(\frac{i}{kr}\right)^2\right]\sin\theta\right\} \tag{2}$$

The magnetic field follows from Faraday's law

$$\overline{H} = \frac{1}{i\omega\mu}\nabla\times\overline{E} = -\hat{\phi}\,ikIl\frac{e^{ikr}}{4\pi r}\left[1+\frac{i}{kr}\right]\sin\theta \tag{3}$$

The fields of the dipole have been extensively studied by Hertz [1893]. Plots of the electric field lines at different times are shown in Figure 4.3.2. The magnetic fields are all in the $\hat{\phi}$ direction circulating the dipole.

The complex Poynting power density is calculated by taking the cross product of $\overline{E}$ and the complex conjugate of $\overline{H}$

$$\overline{S} = \overline{E}\times\overline{H}^*$$
$$= \eta\left[\frac{kIl}{4\pi r}\right]^2\left\{\hat{r}\left[1-\left(\frac{i}{kr}\right)^3\right]\sin^2\theta - \hat{\theta}\left[\left(\frac{i}{kr}\right)-\left(\frac{i}{kr}\right)^3\right]\sin 2\theta\right\} \tag{4}$$

The time-average Poynting power density is

$$<\overline{S}> = \frac{1}{2}Re\{\overline{S}\} = \hat{r}\frac{\eta}{2}\left[\frac{kIl}{4\pi r}\right]^2\sin^2\theta \tag{5}$$

The total radiated power P_r is calculated by integrating $\hat{r}\cdot<\overline{S}>$ over a sphere of radius r with $r\to\infty$. We obtain

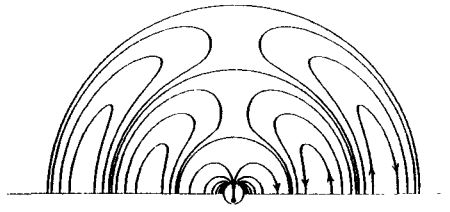
$$P_r = \int_0^{2\pi}d\phi\int_0^{\pi}d\theta\, r^2\sin\theta <S_r> = \frac{4\pi}{3}\eta\left[\frac{kIl}{4\pi}\right]^2 \tag{6}$$

The directive gain $G(\theta,\phi)$ is defined as the power density $S_r(\theta,\phi)$ at observation angles (θ,ϕ) divided by the total radiated power averaged over all angles,

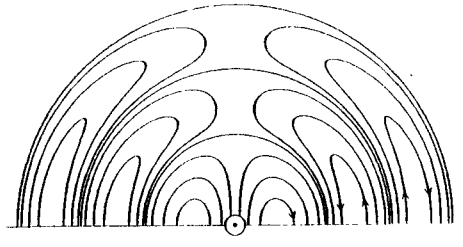
$$G(\theta,\phi) = \frac{<S_r(\theta,\phi)>}{P_r/4\pi r^2} \tag{7}$$

For the Hertzian dipole

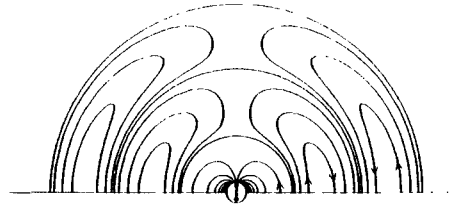
$$G(\theta,\phi) = \frac{3}{2}\sin^2\theta \tag{8}$$



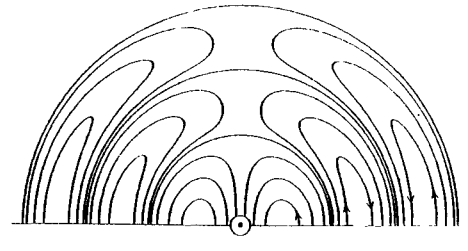
$\omega t = 0$



$\omega t = \pi/2$



$\omega t = \pi$



$\omega t = 3\pi/2$

Figure 4.3.2 Electric field patterns.

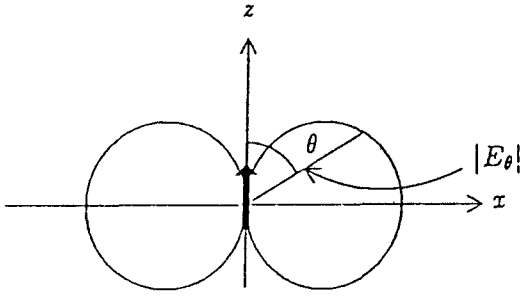


Figure 4.3.3 Radiation field pattern.

The directivity D of an antenna is defined to be the gain at the angle where it is maximum. For the Hertzian dipole,

$$D = G(\theta, \phi)_{max} = \frac{3}{2} \tag{9}$$

which occurs at $\theta = \pi/2$, that is perpendicular to the dipole axis.

When the observation point is very far away from the dipole such that $kr \gg 1$, we can neglect the terms of order higher than $1/kr$ as compared with unity. From (2) and (3), we find the electric and magnetic field in the radiation zone to be

$$\overline{E} = -\hat{\theta}\, i\omega\mu Il \frac{e^{ikr}}{4\pi r} \sin\theta = \hat{\theta} E_\theta \tag{10a}$$

$$\overline{H} = -\hat{\phi}\, ikIl \frac{e^{ikr}}{4\pi r} \sin\theta = \hat{\phi} \frac{1}{\eta} E_\theta \tag{10b}$$

A radiation field pattern can be sketched for the magnitude of $|E_\theta|$ at a constant distance r as a function of the angle θ [Fig. 4.3.3]. The pattern consists of two circles describing $\sin\theta$ and is symmetrical about the z axis. The power pattern or the gain pattern is seen to be proportional to $\sin^2\theta$. As shown in Figure 4.3.4, it is in the form of a horizontal "figure eight" in any plane containing the dipole axis.

b. Hertzian Magnetic Dipole and Small Loop Antenna

The dual of a Hertzian electric dipole is a magnetic dipole with moment ml. Applying the principle of duality, the electromagnetic fields

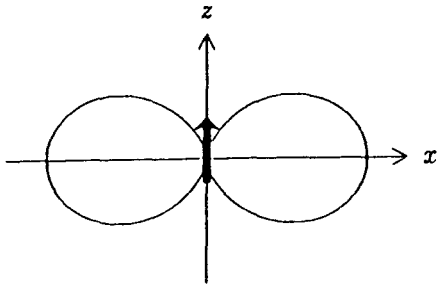


Figure 4.3.4 Radiation power pattern.

are obtained from those for the electric dipole and take the following forms:

$$\overline{E} = \hat{\phi}\, ikml \frac{e^{ikr}}{4\pi r}\left[1 + \frac{i}{kr}\right]\sin\theta$$

$$\overline{H} = -i\omega\epsilon ml \frac{e^{ikr}}{4\pi r}\left\{\hat{r}\left[\frac{i}{kr} + \left(\frac{i}{kr}\right)^2\right] 2\cos\theta + \hat{\theta}\left[1 + \frac{i}{kr} + \left(\frac{i}{kr}\right)^2\right]\sin\theta\right\}$$

A Hertzian magnetic dipole can be realized with the model of a small current loop with radius a and $ml = -i\omega\mu I\pi a^2$ where I is the current in the loop. In the following we discuss the fields produced by such a small current loop.

Consider a small current loop with an infinitesimally small radius a as shown in Figure 4.3.5. Its current density takes the form

$$\overline{J}(\overline{r}') = \hat{\phi} I\delta(\rho' - a)\delta(z') \tag{11}$$

The electric field vector due to the current loop, is calculated from

$$\overline{E}(\overline{r}) = i\omega\mu\left[\overline{\overline{I}} + \frac{1}{k^2}\nabla\nabla\right]\cdot\int_0^{2\pi} d\phi' \int_0^a d\rho' \int_{-\infty}^{\infty} \rho' dz' \frac{\overline{J}(\overline{r}')e^{ik|\overline{r}-\overline{r}'|}}{4\pi\,|\overline{r}-\overline{r}'|} \tag{12}$$

To evaluate the integral, we recognize that in terms of their cartesian components, the radial vectors $\overline{r}$ and $\overline{r}'$ are

$$\overline{r} = \hat{x}\, r\sin\theta\cos\phi + \hat{y}\, r\sin\theta\sin\phi + \hat{z}\, r\cos\theta$$

$$\overline{r}' = \hat{x}\, a\cos\phi' + \hat{y}\, a\sin\phi'$$

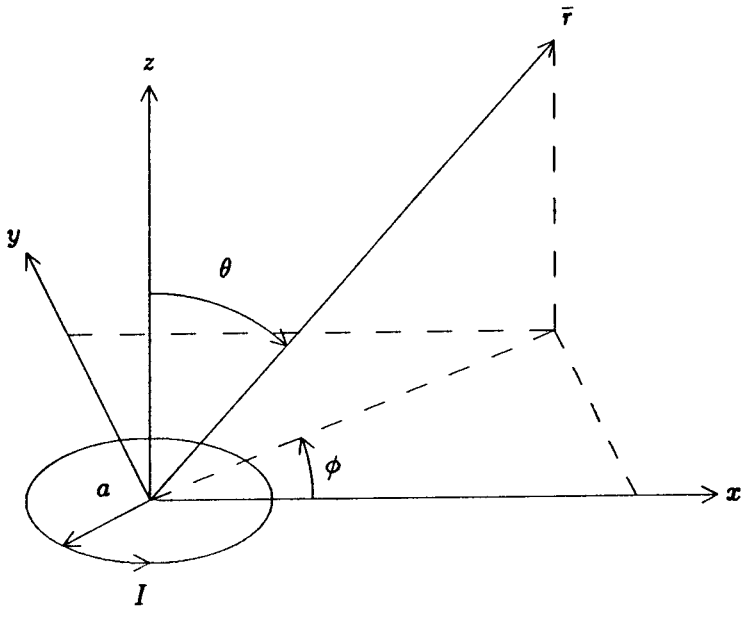


Figure 4.3.5 Small loop antenna.

Because of our choice of the coordinate axes, the loop is on the $x-y$ plane and $\theta' = \pi/2$ for the radial vector $\bar{r}'$. The distance $|\bar{r} - \bar{r}'|$ from observation point $\bar{r}$ to source point $\bar{r}'$ is

$$|\bar{r} - \bar{r}'| = |\hat{x}(r\sin\theta\cos\phi - a\cos\phi') + \hat{y}(r\sin\theta\sin\phi - a\sin\phi') + \hat{z}r\cos\theta|$$

$$= r\sqrt{1 + \frac{a^2}{r^2} - \frac{2a}{r}\sin\theta\cos(\phi - \phi')} \tag{13}$$

We then expand the scalar Green's function in the integral of (2) in the form of a MacLaurin series for $a/r \to 0$. Taking the first two terms, we have

$$\frac{e^{ik|\bar{r}-\bar{r}'|}}{4\pi|\bar{r}-\bar{r}'|} \approx \frac{e^{ikr}}{4\pi r} + \frac{a}{r}\left[\frac{d}{d(\frac{a}{r})}\frac{e^{ik|\bar{r}-\bar{r}'|}}{4\pi|\bar{r}-\bar{r}'|}\right]_{a/r\to 0}$$

$$= \frac{e^{ikr}}{4\pi r} + \frac{a}{r}(-ikr + 1)\sin\theta\cos(\phi - \phi')\frac{e^{ikr}}{4\pi r} \tag{14}$$

Noticing that $\hat{\phi} = -\hat{x}\sin\phi' + \hat{y}\cos\phi'$, we evaluate the integral in (12) by substituting (14)

$$\int_0^{2\pi} a\,d\phi'(-\hat{x}\sin\phi' + \hat{y}\cos\phi')\frac{Ie^{ik|\bar{r}-\bar{r}'|}}{4\pi|\bar{r}-\bar{r}'|}$$

$$= (-\hat{x}\sin\phi + \hat{y}\cos\phi)\frac{\pi a^2 I e^{ikr}}{4\pi r^2}(1 - ikr)\sin\theta$$

$$= \hat{\phi}\frac{I\pi a^2 e^{ikr}}{4\pi r^2}(1 - ikr)\sin\theta \tag{15}$$

Substituting the above result back into (12) and noticing that (15) is independent of ϕ, we see that the del operator $\nabla\nabla$ does not contribute. The electric field vector becomes

$$\overline{E} = \hat{\phi}\omega\mu k I\pi a^2 \frac{e^{ikr}}{4\pi r}\left[1 + \frac{i}{kr}\right]\sin\theta \tag{16}$$

The magnetic field vector is

$$\overline{H}(\overline{r}) = \frac{1}{i\omega\mu}\nabla \times \overline{E}(\overline{r})$$

$$= -k^2 I\pi a^2 \frac{e^{ikr}}{4\pi r}\left\{\hat{r}\left[\frac{i}{kr} + \left(\frac{i}{kr}\right)^2\right]2\cos\theta + \hat{\theta}\left[1 + \frac{i}{kr} + \left(\frac{i}{kr}\right)^2\right]\sin\theta\right\} \tag{17}$$

The electric and magnetic fields are seen to be the dual of those for the Hertzian dipole. Outside the current loop, the magnetic field pattern is exactly the same as the electric field lines of the Hertzian dipole. Discussion on the Poynting power vector, the gain, and the radiation fields is also exactly the same for the two elementary radiating elements.

The correspondence can be quantified by letting the Hertzian dipole moment Il be

$$(Il)_e = (ikI\pi a^2)_m \tag{18}$$

where the subscripts e and m are used to denote the electric dipole and the current loop, respectively. The solution for a small current loop in (16) and (17) can be obtained from the solutions for electric dipoles by letting

$$\overline{E}_e = \eta\overline{H}_m \tag{19}$$

$$\overline{H}_e = -\frac{\overline{E}_m}{\eta} \tag{20}$$

Such relations as illustrated in (10)–(20) depict the duality principle. What we have shown is that the complete solution for a small current loop, including both near and far fields, is the dual of that for

a Hertzian electric dipole. A small loop can thus be regarded as a magnetic dipole that is the dual of an electric dipole.

4.4 Radiation Fields

When the observation point is very far away from the source, approximations can be made in the evaluation of the radiation or far fields. The radiation field approximation consists of the following two conditions:

$$|\bar{r} - \bar{r}'| \approx r - \hat{r} \cdot \bar{r}' \tag{1}$$

$$kr \gg 1 \tag{2}$$

From Figure 4.4.1 we see that the line joining the remote observation point to the origin is almost parallel to the line connecting the observation point to the source points where integration is performed. In the radiation zone, the $\bar{k}$ vector is in the $\hat{r}$ direction, $\bar{k} = \hat{r}k$. We find

$$\begin{aligned} \bar{E}(\bar{r}) &= i\omega\mu \left[\bar{\bar{I}} + \frac{1}{k^2}\nabla\nabla\right] \cdot \iiint dV' \, \frac{\bar{J}(\bar{r}')e^{ik|\bar{r}-\bar{r}'|}}{4\pi\,|\bar{r}-\bar{r}'|} \\ &\approx i\omega\mu \left[\bar{\bar{I}} + \frac{1}{k^2}\nabla\nabla\right] \cdot \frac{e^{ikr}}{4\pi r} \iiint dV' \, \bar{J}(\bar{r}')e^{-i\bar{k}\cdot\bar{r}'} \end{aligned} \tag{3}$$

In the approximation, we neglect $\hat{r} \cdot \bar{r}'$ in the denominator. The term $k\hat{r} \cdot \bar{r}'$ is kept in the exponent in (3) because its contribution to the phase variation can be significant when it is of the order of, or larger than, π.

We define a vector current moment

$$\bar{f}(\theta, \phi) = \iiint dV' \, \bar{J}(\bar{r}')e^{-i\bar{k}\cdot\bar{r}'} \tag{4}$$

We see that the current density $\bar{J}(\bar{r}')$, weighted with the phase-retardation factor $e^{-i\bar{k}\cdot\bar{r}'}$, is integrated over the volume. Since the integrand is a function of $\bar{r}'$, the current moment after integration will be a function of θ and ϕ only and independent of the observation distance from the origin r.

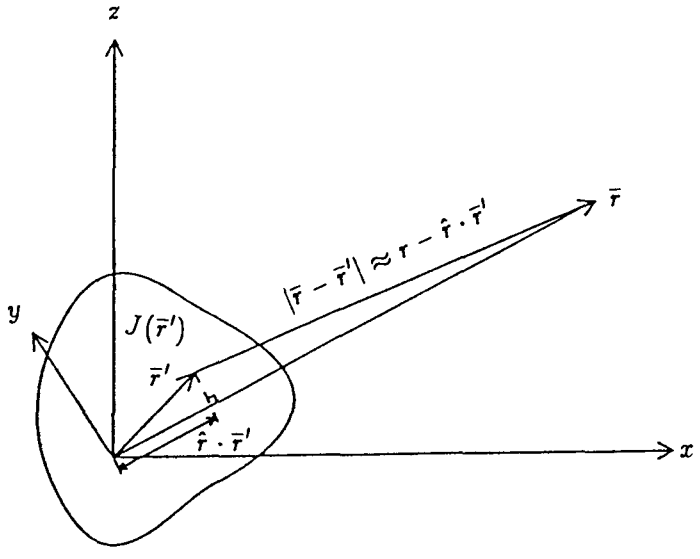


Figure 4.4.1 Far-field approximation.

The del operator ∇ in (3) can be replaced by $i\bar{k}$ in the far-field approximation. For instance, consider the gradient

$$\nabla = \hat{r}\frac{\partial}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial}{\partial \theta} + \hat{\phi}\frac{1}{r\sin\theta}\frac{\partial}{\partial \phi}$$

The operator $\partial/\partial r$ when operated on e^{ikr} gives ik which yields a term of the order of $1/r$. All other terms of the del operator give rise to terms of the order of $(1/r)^2$ or higher. Under the far field approximation of $kr \gg 1$, we keep only the term of the order of $1/r$ and replace the del operator by $i\bar{k}$. The radiated electric field becomes

$$\bar{E}(\bar{r}) = i\omega\mu[\bar{\bar{I}} - \hat{r}\hat{r}] \cdot \bar{f}\frac{e^{ikr}}{4\pi r} = i\omega\mu\frac{e^{ikr}}{4\pi r}(\hat{\theta}f_\theta + \hat{\phi}f_\phi) \tag{5}$$

The term $\bar{f}e^{ikr}/4\pi r$ is also referred to as the radiation vector. The magnetic field $\bar{H}(\bar{r})$ is, under the same far-field approximation,

$$\bar{H}(\bar{r}) = \frac{1}{i\omega\mu}\nabla \times \bar{E}(\bar{r}) = \frac{\bar{k}}{\omega\mu} \times \bar{E}(\bar{r}) = \frac{ike^{ikr}}{4\pi r}(\hat{\phi}f_\theta - \hat{\theta}f_\phi) \tag{6}$$

The time-average Poynting's power density is

$$< \overline{S} > = \frac{1}{2} Re \left\{ \overline{E} \times \overline{H}^* \right\} = \hat{r} \frac{1}{2} \sqrt{\frac{\mu}{\epsilon}} \left(\frac{k}{4\pi r} \right)^2 (|f_\theta|^2 + |f_\phi|^2) \quad (7)$$

Thus, to calculate the radiation field for a given source $\overline{J}$, the first task is to evaluate the vector current moment $\overline{f}(\theta, \phi)$.

For the Hertzian dipole with $\overline{J}(\vec{r}') = \hat{z} Il \, \delta(\vec{r}')$, we find from (4) the vector current moment

$$\overline{f}(\theta, \phi) = \hat{z} Il = (\hat{r} \cos\theta - \hat{\theta} \sin\theta) Il$$

From (5) to (7), we find that the electric field, the magnetic field, and the time-average Poynting power density are identical to those for the radiation field of the Hertzian electric dipole.

a. Radiation Resistance

From the input terminals of an antenna, the field radiated by the antenna can be viewed as power dissipated in a resistor called the radiation resistance of the antenna. For a wire antenna with a current distribution $\overline{J}(\vec{r}') = \hat{z} I(z') \delta(x') \delta(y')$, the θ component of the vector current moment is

$$f_\theta = -\sin\theta \iiint d^3 r' \, J(\vec{r}') e^{-i\vec{k} \cdot \vec{r}'} = -\sin\theta \int dz' \, I(z') e^{-ikz' \cos\theta} \quad (8)$$

The electric field is calculated as

$$E_\theta = iw\mu \frac{e^{ikr}}{4\pi r} f_\theta$$

and the magnetic field as

$$H_\phi = ik \frac{e^{ikr}}{4\pi r} f_\theta$$

The total radiated power is

$$\begin{aligned} P_r &= \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \, r^2 \sin\theta \, \frac{1}{2} E_\theta H_\phi^* \\ &= \int_0^{\pi} d\theta \sin\theta \, \pi\eta \left(\frac{k}{4\pi} \right)^2 |f_\theta|^2 \end{aligned} \quad (9)$$

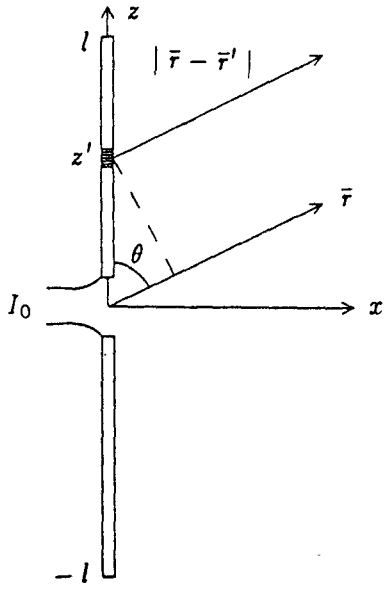


Figure 4.4.2 Wire antenna excited by a current source.

where $\eta = (\mu/\epsilon)^{1/2}$. The radiation resistance is

$$R_r = 2P_r/I_0^2$$

where I_0 is the input current amplitude.

Consider a wire antenna with zero radius and length $2l$ excited at the center bv a constant current source I_0 [Fig. 4.4.2]. When l is infinitesimally small, we may use the Hertzian dipole model with $I(z') = 2I_0 l\delta(z')$. The current moment is found from (8) to be $f_\theta = -2I_0 l \sin\theta$. The total radiated power for a wire antenna with the current moment f_θ is calculated from (9) to be

$$P_r = \int_0^\pi d\theta \sin^3\theta\,(k/4\pi)^2 4\pi\eta\, I_0^2 l^2 = \frac{(2kI_0 l)^2}{12\pi}\eta$$

Thus the radiation resistance

$$R_r = \frac{2P_r}{I_0^2} = \frac{(2kl)^2}{6\pi}\eta = 20(2kl)^2 \qquad (10)$$

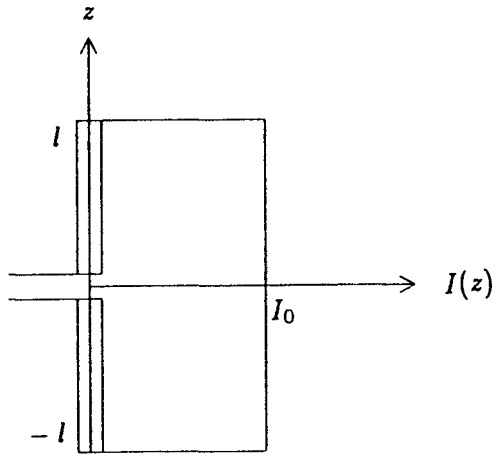


Figure 4.4.3 Uniform current distribution.

where we made use of the fact that $\eta = 120\pi$ ohms. Clearly, the radiation resistance R_r as calculated in (10) is not applicable to a wire antenna with any significant length. For instance, when $l = \lambda/2$, we will find $R_r \approx 200$ ohms where the more accurate value from measurement and from more rigorous calculation is 73 ohms.

In the calculation of R_r in (10), we assume the current distribution on the wire is uniform [Fig. 4.4.3]. If the wire is made of conductors, boundary conditions at $z = \pm l$ will require $I(z)$ be zero there. A better approximation for the current distribution for a small l takes the form of a triangular shape [Fig. 4.4.4],

$$I(z') = \frac{I_0}{l}(l - |z'|)$$

In the calculation of the current moment, we first neglect the phase retardation factor $e^{-ikz'\cos\theta}$ as the wire length is small,

$$f_\theta \approx -\sin\theta \int_{-l}^{l} dz' \frac{I_0}{l}(l - |z'|) = -I_0 l \sin\theta \tag{11}$$

This is equivalent to replacing the current moment $-2I_0 l \sin\theta$ for a Hertzian dipole of uniform current distribution as shown in Figure 4.4.3 with the average current moment $-I_0 l \sin\theta$ of the triangular distributions as shown in Figure 4.4.4. The radiation resistance is calculated to be

$$R_r = \frac{2P_r}{I_0^2} = 5(2kl)^2$$

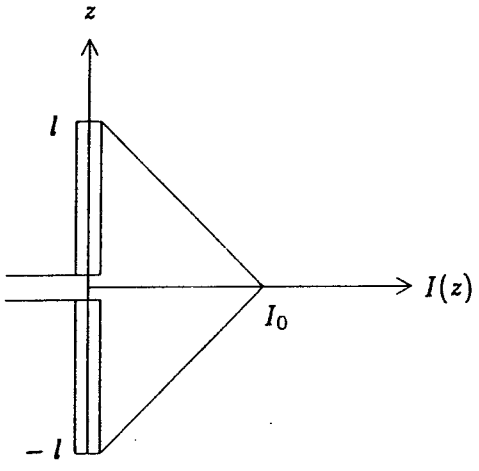


Figure 4.4.4 Triangular current distribution.

which is four times smaller than that in (10). Stretching the limit of $2l$ to a half-wavelength, we find $R_r = 50$ ohms. For a wire antenna with $2l$ smaller than $\lambda/2$, taking into account the phase-retardation factor will serve to decrease the current moment because the integrand in (11) is always positive and the phase-retardation factor causes its magnitude to decrease. Therefore the reason for the underestimation of R_r cannot be attributed to the omission of the phase-retardation factor. It must be due to the inaccuracy in the assumption of the current distribution.

For a wire antenna that is not infinitesimally small, the more accurate current distribution is approximated by a sinusoidal distribution. We assume [Fig. 4.4.5]

$$I(z') = I_m \sin k(l - |z'|) \tag{12}$$

Notice that I_m is the amplitude of the sinusoid that is related to I_0 by $I_0 = I_m \sin(kl)$. The θ component of the vector current moment is

$$f_\theta = -\sin\theta \int_{-l}^{l} dz'\, I_m \sin k(l - |z'|) e^{-ikz' \cos\theta}$$

The integration can be evaluated and gives rise to a closed-form formula.

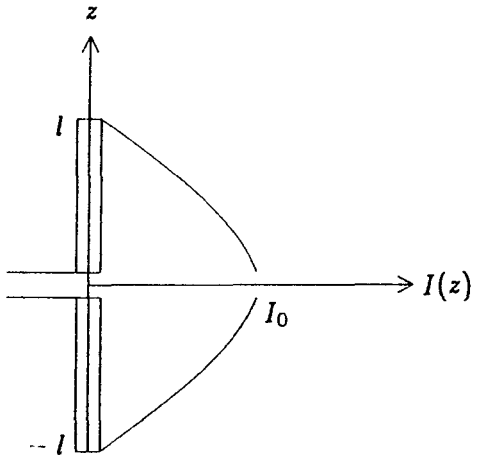


Figure 4.4.5 Sinusoidal current distribution.

Before evaluating the integral exactly we again assume that the length $2l$ is small and that the contribution from the phase factor $e^{-ikz'\cos\theta}$ is uniformly close to 1. Evaluating the integral without taking into account the phase factor, we find

$$f_\theta = -\sin\theta\frac{2I_m}{k}(1-\cos kl) \tag{13}$$

Under the limit of $kl \ll 1$, we get the approximate result

$$f_\theta \approx -I_m(kl^2)\sin\theta \tag{14}$$

Notice that $I_m \approx I_0/kl$. Thus (14) is identical to (11).

Applying the result to a half-wavelength wire, we let $I_m = I_0$ and obtain from (13), $f_\theta = -\sin\theta(2I_0/k)$. The total radiated power is calculated to be, by using (9)

$$P_r = \int_0^\pi d\theta \sin^3\theta \frac{\eta I_0^2}{4\pi} = 40I_0^2 \tag{15}$$

where $\eta = \sqrt{\mu/\epsilon} = 120\pi$ ohms. It follows that the radiation resistance is $R_r = 80$ ohms. The reason that the sinusoidal current distribution gives a larger radiation resistance than the triangular assumption is

that the area under the sine is larger then the triangle. However, the phase-retardation effect has been neglected. If it is not neglected, the result will be applicable to wire antennas with any length, small or large.

The current moment for a wire antenna of length $2l$ and with the sinusoidal current distribution, taking into account the phase-retardation factor, is calculated as

$$f_\theta = -\sin\theta \int_{-l}^{l} dz'\, I_m \sin k(l - |z'|) e^{-ikz' \cos\theta}$$
$$= -\frac{2I_m}{k}\left[\frac{\cos(kl\cos\theta) - \cos kl}{\sin\theta}\right]$$

The electric field vector is

$$\overline{E} = \hat{\theta}\, i\omega\mu \frac{e^{ikr}}{4\pi r} f_\theta = -\hat{\theta}\frac{i\eta I_m e^{ikr}}{2\pi r}\left[\frac{\cos(kl\cos\theta) - \cos kl}{\sin\theta}\right]$$

The total radiated power P_r is

$$P_r = \int_0^\pi d\theta\, \sin\theta \frac{\eta {I_m}^2}{4\pi}\left[\frac{\cos(kl\cos\theta) - \cos kl}{\sin\theta}\right]^2$$

The radiation resistance is, using I_m as the input current

$$R_r = \frac{2P_r}{{I_m}^2} = \frac{\eta}{2\pi}\int_0^{2\pi} d\theta \frac{[\cos(kl\cos\theta) - \cos kl]^2}{\sin\theta \sin^2 kl}$$

The radiation resistance R_r can be evaluated in terms of tabulated functions. The result is

$$R_r = \frac{\eta}{2\pi}\left\{\gamma + \ln(2kl) - Ci(2kl) + \sin(2kl)\left[\frac{1}{2}Si(4kl) - Si(2kl)\right]\right.$$
$$\left. + \frac{1}{2}\cos(2kl)\left[\gamma + \ln(kl) + Ci(4kl) - 2Ci(2kl)\right]\right\}$$

where

$$Si(x) = \int_0^x dx' \frac{\sin x'}{x'}$$
$$Ci(x) = -\int_x^\infty dx' \frac{\cos x'}{x'}$$

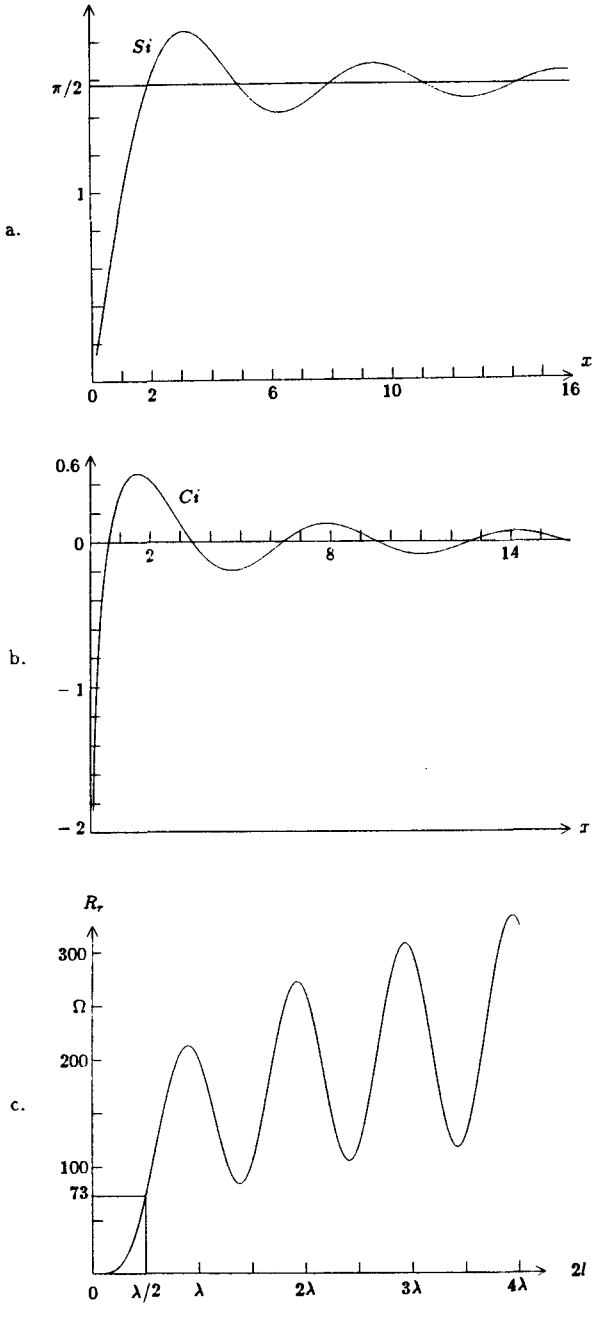


Figure 4.4.6 Functions $Si(x)$ and $Ci(x)$ and radiation resistance R_r.

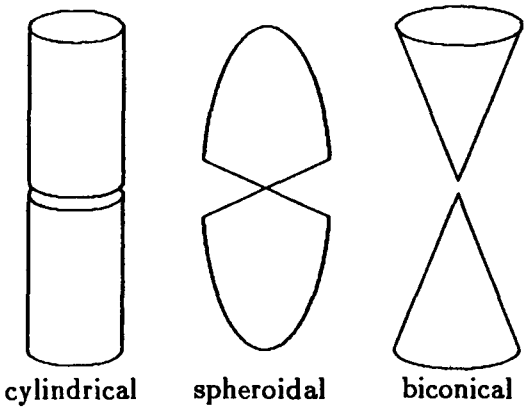


Figure 4.4.7 Models of antenna.

and $\gamma = 0.5772\ldots$ The sine integral $Si(x)$ and the cosine integral $Ci(x)$ are illustrated in Figures 4.4.6a,b. A numerical plot for R_r as a function of $2l$ is shown in Figure 4.4.6c. When $2l = \lambda/2$, we find $R_r = 73$ ohms.

In the above discussions, we have illustrated how an antenna problem is attacked. The results clearly become more accurate as the assumptions become more realistic, which also bring in more complicated mathematical manipulations. Nevertheless, we cannot claim that the wire antenna problem has now acquired a satisfactory answer. Two major problems still remain: (i) The radius of the wire has been tacitly assumed to be infinitesimally small. What if it is not? (ii) How do we know whether the sinusoidal current distribution on the wire is the valid assumption? Why do we not assume other types of current distributions which also satisfy the boundary conditions of zero current at $z = \pm l$? These questions can be answered only with a detailed model of the antenna. With a precise physical structure, the current on the antenna can be *determined* from Maxwell's equations, *not assumed*.

To address the above questions, at least three different models [Fig. 4.4.7] have been used with a rigorous boundary value problem approach. The model by Hallen and King [1969] is a perfectly conducting cylinder with finite radius a and length l with a gap at $z = 0$ where the electric field strength is $-V\delta(z)$. The problem is formulated in the form of integral equations. The second approach is due to Stratton and Chu [1941] who modelled a dipole antenna as a pair of ellipsoids

or prolate spheroids. Normal modes associated with the structure are used to match boundary conditions. The third approach is developed by Schelkunoff [1943] with the model of a biconical structure. In the next section we shall pursue Schelkunoff's approach to the solution of this fundamental problem.

4.5 Biconical Antennas

a. Formulation and Wave Solutions

The geometrical configuration of a biconical antenna is shown in Figure 4.5.1. Due to the symmetry, we expect that all field solutions are ϕ-independent. Maxwell's equations can then be reduced to two uncoupled sets of equations; one set relates E_r, E_θ, and H_ϕ and forms the TM modes, and the other set relates H_r, H_θ, and E_ϕ and forms the TE modes. For the biconical antenna problem, the currents are in the $\hat{r}$ direction and we expect TM waves solutions.

We deal with the solutions of Maxwell's equations satisfying the following boundary conditions:

(i) In the air region, the electromagnetic field must be an outgoing wave.

(ii) In the antenna region, the tangential electric field vanishes on the surfaces of the cones

$$E_r(\theta_0) = E_r(\pi - \theta_0) = 0 \tag{1}$$

At the input,

$$V_0 = \lim_{r\to 0} \int_{\theta_0}^{\pi-\theta_0} d\theta\, r E_\theta \tag{2}$$

where V_0 is the given external excitation voltage.

(iii) On the boundary surface S_0 separating the antenna region and the air region

$$E_\theta(l+0) = 0 \qquad 0 \le \theta \le \theta_0 \text{ and } \pi - \theta_0 \le \theta \le \pi \tag{3}$$

$$E_\theta(l+0) = E_\theta(l-0) \qquad \theta_0 < \theta < \pi - \theta_0 \tag{4}$$

$$E_r(l+0) = E_r(l-0) \qquad \theta_0 < \theta < \pi - \theta_0 \tag{5a}$$

$$H_\phi(l+0) = H_\phi(l-0) \qquad \theta_0 < \theta < \pi - \theta_0 \tag{5b}$$

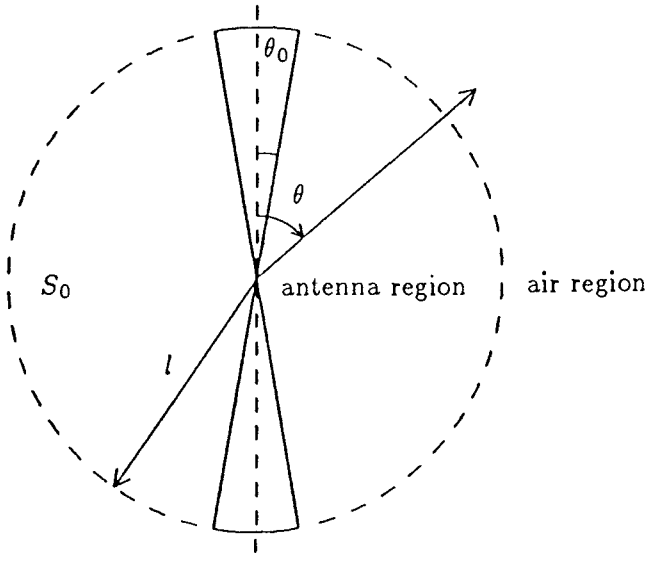


Figure 4.5.1 Biconical antenna.

Note that boundary conditions (5*a*) and (5*b*) are not independent of each other because (5*a*) is associated with Gauss's law and (5*b*) with Ampere's law. They depend on each other as a consequence of the charge–current conservation law. We shall use either (5*a*) or (5*b*), depending on whichever is more convenient. The governing equations for the TM waves are

$$\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\, H_\phi) = -i\omega\epsilon E_r \tag{6a}$$

$$-\frac{1}{r}\frac{\partial}{\partial r}(r\, H_\phi) = -i\omega\epsilon E_\theta \tag{6b}$$

$$\frac{1}{r}\frac{\partial}{\partial r}(rE_\theta) - \frac{1}{r}\frac{\partial}{\partial\theta}(E_r) = i\omega\mu H_\phi \tag{6c}$$

Substituting (6*a*) and (6*b*) in (6*c*) to eliminate E_r and E_θ, we obtain

$$\frac{\partial^2}{\partial r^2} r H_\phi + \frac{1}{r^2}\frac{\partial}{\partial\theta}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\, r\, H_\phi)\right] + k^2 r\, H_\phi = 0 \tag{7}$$

Equation (7) is solved by the separation of variables technique. Let

$$H_\phi = R(r)\Theta(\theta) \tag{8}$$

We obtain the following two equations

$$\frac{1}{r}\frac{d^2}{dr^2}rR(r) + \left[k^2 - \frac{n(n+1)}{r^2}\right]R(r) = 0 \tag{9}$$

$$\frac{d}{d\theta}\left[\frac{1}{\sin\theta}\frac{d}{d\theta}(\sin\theta\Theta(\theta))\right] + n(n+1)\Theta(\theta) = 0 \tag{10}$$

with $n(n+1)$ as the separation constant.

Solutions to (9) are the spherical Hankel functions $h_n^{(1)}(kr)$ and $h_n^{(2)}(kr)$. For the first few orders $n = 0$ to $n = 2$, the spherical Hankel functions of first kind are

$$h_0^{(1)}(kr) = -i\frac{e^{ikr}}{kr} \tag{11a}$$

$$h_1^{(1)}(kr) = -\left[1 + \frac{i}{kr}\right]\frac{e^{ikr}}{kr} \tag{11b}$$

$$h_2^{(1)}(kr) = i\left[1 + 3\left(\frac{i}{kr}\right) + 3\left(\frac{i}{kr}\right)^2\right]\frac{e^{ikr}}{kr} \tag{11c}$$

The second kind spherical Hankel functions $h_n^{(2)}(kr)$ are complex conjugates of $h_n^{(1)}(kr)$. The spherical Bessel functions $b_n(\xi)$ are related to the cylindrical Bessel functions $B_n(\xi)$ by

$$b_n(\xi) = \sqrt{\frac{\pi}{2\xi}}B_{n+\frac{1}{2}}(\xi)$$

with $b_n(\xi)$ representing $h_n^{(1)}$ or $h_n^{(2)}$ and $B_n(\xi)$ representing $H_n^{(1)}$ or $H_n^{(2)}$. The recurrence formulas are

$$\begin{aligned} B'_\nu(\xi) &= B_{\nu-1}(\xi) - \frac{\nu}{\xi}B_\nu(\xi) \\ &= -B_{\nu+1}(\xi) + \frac{\nu}{\xi}B_\nu(\xi) \end{aligned}$$

The spherical Bessel functions $j_n(kr)$ are defined as the real part of either $h_n^{(1)}(kr)$ or $h_n^{(2)}(kr)$. For the first few orders, they are

$$j_0(kr) = \frac{\sin kr}{kr} \tag{12a}$$

$$j_1(kr) = -\frac{\cos kr}{kr} + \frac{\sin kr}{(kr)^2} \tag{12b}$$

$$j_2(kr) = -\frac{\sin kr}{kr} - \frac{3\cos kr}{(kr)^2} + \frac{3\sin kr}{(kr)^3} \tag{12c}$$

Note that the $j_n(kr)$ functions remain finite as $kr \to 0$. The imaginary parts of $h_n^{(1)}(kr)$ are known as the spherical Neumann functions, which become infinite as $kr \to 0$.

Solutions to (10) are derivatives of the Legendre polynomial $P(\theta)$

$$\Theta(\theta) = \frac{d}{d\theta} P(\theta) \tag{13}$$

where $P(\theta)$ satisfies the ordinary Legendre equation

$$\frac{1}{\sin\theta} \frac{d}{d\theta} \left[\sin\theta \frac{dP(\theta)}{d\theta} \right] + n(n+1)P(\theta) = 0 \tag{14}$$

The two independent solutions to (14) are $P_n(\cos\theta)$ and $Q_n(\cos\theta)$ where

$$P_n(\cos\theta) = \sum_{q=0}^{\infty} \frac{(-1)^q (n+q)!}{(n-q)!(q!)^2} \sin^{2q}(\theta/2) \tag{15}$$

The other independent solution is $Q_n(\cos\theta) = P_n(-\cos\theta)$ for noninteger values of n. For integer values of n, we have

$$Q_n(\cos\theta) = P_n(\cos\theta) \ln\left(\cot\frac{\theta}{2}\right) - \sum_{s=1}^{n} \frac{P_{n-s} P_{s-1}}{s} \tag{16}$$

where the summation term contributes when $n \geq 1$. The Rodrigues' formula for generating Legendre polynomials states, for integer values of n,

$$P_n(u) = \frac{1}{2^n n!} \frac{d^n}{du^n} (u^2 - 1)^n$$

where $u = \cos\theta$. The first few Legendre polynomials $P_n(\cos\theta)$ are

$$P_0(\cos\theta) = 1 \tag{17a}$$

$$P_1(\cos\theta) = \cos\theta \tag{17b}$$

$$P_2(\cos\theta) = \frac{1}{2}(3\cos^2\theta - 1) = \frac{1}{4}(3\cos 2\theta + 1) \tag{17c}$$

Note that the $Q_n(\cos\theta)$ functions for integer values of n become infinite at $\theta = 0$ and $\theta = \pi$. For noninteger values of n the functions $P_n(-\cos\theta)$ become infinite at $\theta = \pi$.

In terms of $R(r)$ and $P(\theta)$, we now find from (6a), (6b), and (8)

$$H_\phi = R(r)\frac{dP(\theta)}{d\theta} \tag{18a}$$

$$E_r = -\frac{R(r)}{i\omega\epsilon r}\frac{1}{\sin\theta}\frac{d}{d\theta}\left[\sin\theta\frac{dP(\theta)}{d\theta}\right] = \frac{n(n+1)}{i\omega\epsilon r}R(r)P(\theta) \tag{18b}$$

$$E_\theta = \frac{1}{i\omega\epsilon r}\frac{d[rR(r)]}{dr}\frac{dP(\theta)}{d\theta} \tag{18c}$$

b. Solution in the Air Region and Dipole Fields

We first consider solutions in the air region. They take the form, for $n = 0, 1, 2, \ldots$

$$H_\phi = \frac{1}{2\pi}\sum_n b_n h_n^{(1)}(kr)\frac{d}{d\theta}P_n(\cos\theta) \tag{19a}$$

$$E_r = \frac{1}{2\pi i\omega\epsilon r}\sum_n n(n+1)b_n h_n^{(1)}(kr)P_n(\cos\theta) \tag{19b}$$

$$E_\theta = \frac{1}{2\pi i\omega\epsilon r}\sum_n b_n\frac{d}{dr}[rh_n^{(1)}(kr)]\frac{d}{d\theta}P_n(\cos\theta) \tag{19c}$$

We choose the Hankel functions of the first kind because they represent outgoing waves. The functions with noninteger values of n and the $Q_n(\cos\theta)$ solutions are excluded because they are singular in the directions $\theta = 0$ and $\theta = \pi$. The electric field lines for the TM_1 and TM_2 modes are plotted in Figures 4.5.2–3

We note that the TEM (or the TM_0) mode is not present in the air region because $n = 0$ and $P_0(\cos\theta) = 1$. For the TM_1 mode, we find

$$H_\phi = \frac{b_1 e^{ikr}}{2\pi kr}\left[1 + \frac{i}{kr}\right]\sin\theta \tag{20a}$$

$$E_r = \frac{\eta b_1 e^{ikr}}{\pi kr}\left[\frac{i}{kr} + \left(\frac{i}{kr}\right)^2\right]\cos\theta \tag{20b}$$

$$E_\theta = \frac{\eta b_1 e^{ikr}}{2\pi kr}\left[1 + \frac{i}{kr} + \left(\frac{i}{kr}\right)^2\right]\sin\theta \tag{20c}$$

where $\eta = k/\omega\epsilon = (\mu/\epsilon)^{1/2}$. These are the fields for a Hertzian dipole. We can determine b_1 by using the cylindrical coordinate system and

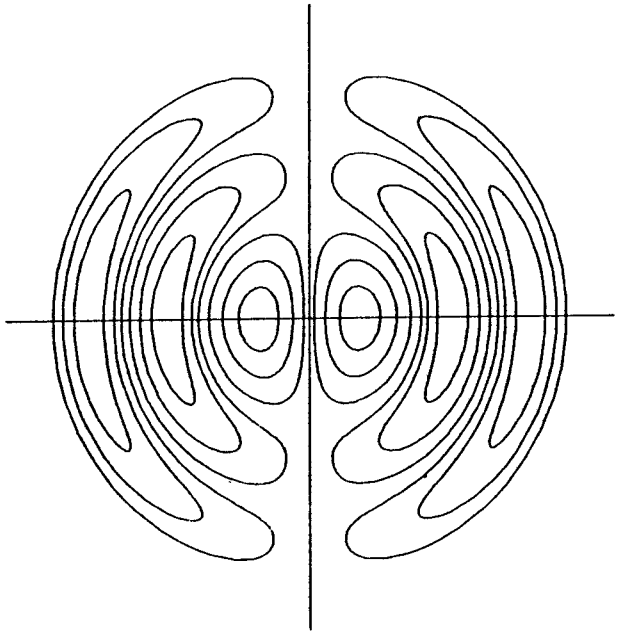


Figure 4.5.2 Electric field lines for TM_1 mode.

noting that the current dipole moment is

$$I_0 l_{eff} = \int_{-\infty}^{\infty} dz\, I_z \tag{21}$$

where

$$I_z = \int_0^{2\pi} \rho\, d\phi\, H_\phi$$

As $r = (z^2 + \rho^2)^{1/2}$ and $\sin\theta = \rho/r$, we obtain from (20a) as $\rho \to 0$ and $z \to 0$

$$I_z \approx \frac{ib_1\rho^2}{k^2(z^2+\rho^2)^{3/2}}$$

Notice that $I_z \to 0$ as $\rho \to 0$. At the point $z = 0$, however, $I_z \to \infty$ as $\rho \to 0$. Carrying out the integral (21), we find

$$I_0 l_{eff} = \frac{i2b_1}{k^2}\int_0^{\infty} dz \frac{\rho^2}{(z^2+\rho^2)^{3/2}} = \frac{i2b_1}{k^2}$$

which yields

$$b_1 = \frac{k^2}{i2} I_0 l_{eff}$$

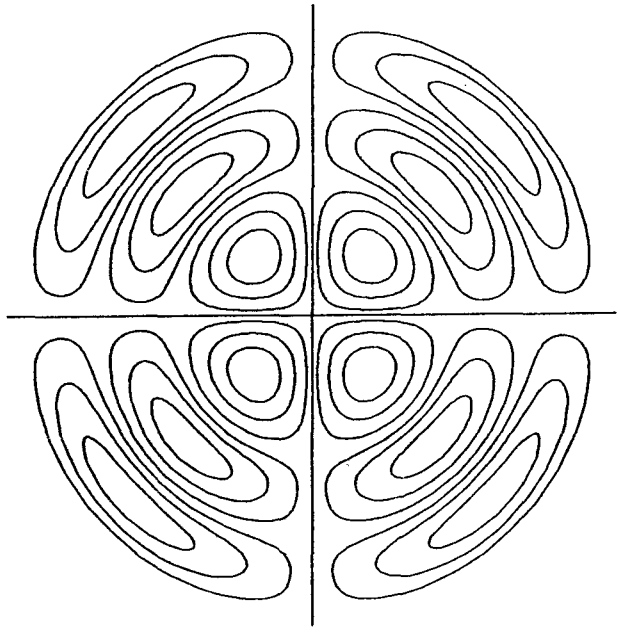


Figure 4.5.3 Electric field lines for TM_2 mode.

The solutions expressed in (20) are identical to those for an infinitesimal dipole antenna. The electric field lines for the quadrupole field is illustrated in Figure 4.5.3.

c. Solution in the Antenna Region

In the antenna region, we note from (18*b*) that the boundary conditions for E_r as expressed in (1) are satisfied either when $n = 0$ or when

$$P(\theta_0) = P(\pi - \theta_0) = 0 \tag{22}$$

When $n = 0$, we have the TEM mode. The solution for $R(r)$ is a linear combination of e^{ikr}/r and e^{-ikr}/r, which represent waves guided by the cones and reflected at the terminals. The $Q_0(\cos\theta)$ function is used since $P_0(\cos\theta) = 1$ and $d[P_0(\cos\theta)]/d\theta = 0$.

For higher-order TM modes, n is determined from (22) and is in general not an integer. We denote it with u. We choose $P(\theta)$ to be a linear combination of $P_u(\cos\theta)$ and $P_u(-\cos\theta)$:

$$P(\theta) = T_u(\theta) = \frac{1}{2}[P_u(\cos\theta) + aP_u(-cos\theta)]$$

Equation (22) requires that

$$P_u(\cos\theta_0) + aP_u(-\cos\theta_0) = 0$$
$$P_u(-\cos\theta_0) + aP_u(\cos\theta_0) = 0$$

We find that $a = \pm 1$. Therefore we have either

$$T_u(\theta) = \frac{1}{2}\left[P_u(\cos\theta) - P_u(-\cos\theta)\right]$$

which is an odd function in $\cos\theta$, or

$$T_u(\theta) = \frac{1}{2}\left[P_u(\cos\theta) + P_u(-\cos\theta)\right]$$

which is an even function. The derivatives

$$\frac{d}{d\theta}\left\{\frac{1}{2}\left[P_u(\cos\theta) - P_u(-\cos\theta)\right]\right\} = -\frac{1}{2}\left[P_u'(\cos\theta) + P_u'(-\cos\theta)\right]\sin\theta$$
$$\frac{d}{d\theta}\left\{\frac{1}{2}\left[P_u(\cos\theta) + P_u(-\cos\theta)\right]\right\} = -\frac{1}{2}\left[P_u'(\cos\theta) - P_u'(-\cos\theta)\right]\sin\theta$$

are, respectively, even and odd functions. Hence H_ϕ due to odd $T_u(\theta)$ functions are even functions of θ, which means that the currents in the upper and the lower cones are in the same direction. (If current at r_0 in the upper cone is flowing away from the origin, then the current at r_0 in the lower cone is flowing toward the origin.) For the even $T_u(\theta)$ functions, H_ϕ is an odd function of θ, which means the current in the upper and lower cones at some distance r_0 will be flowing in opposite directions. (Both currents will be flowing away from or toward the origin.)

We will only consider the balanced type of feed described by a series of odd $T_u(\theta)$ functions, which is the most important case from the practical point of view. Thus we choose

$$T_u(\theta) = \frac{1}{2}\left[P_u(\cos\theta) - P_u(-\cos\theta)\right] \tag{23}$$

The complete solution may be written as a summation over u with u indicating the nearest integer of u,

$$H_\phi = \frac{I_0(r)}{2\pi r\sin\theta} + \frac{1}{2\pi}\sum_u a_u j_u(kr)\frac{d}{d\theta}T_u(\theta) \tag{24a}$$

$$E_r = \frac{1}{2\pi r i\omega\epsilon}\sum_u u(u+1)a_u j_u(kr)T_u(\theta) \tag{24b}$$

$$E_\theta = \frac{1}{2\pi r i\omega\epsilon\sin\theta}\frac{d}{dr}I_0(r) + \frac{1}{2\pi r i\omega\epsilon}\sum_u a_u\frac{d}{dr}\left[rj_u(kr)\right]\frac{d}{d\theta}T_u(\theta) \tag{24c}$$

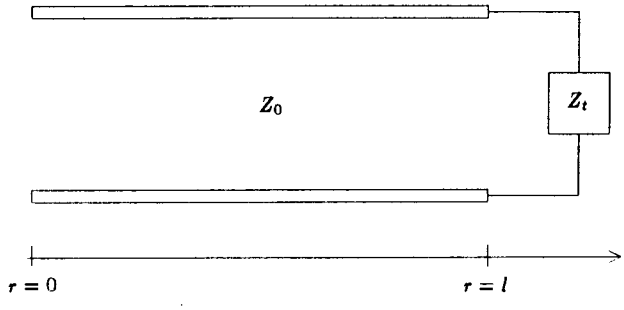


Figure 4.5.4 Transmission line model for the TEM mode.

where $T_u(\theta)$ is given in (23). The first terms in (24*a*) and (24*c*) are the TEM solutions obtained from (9) and (10) with $n = 0$ and $I_0(r) \sim e^{\pm ikr}$. For higher-order modes, we choose the spherical Bessel functions $j_u(kr)$ only. This is because when the Neumann function $N_u(kr)$ is included, not only would the field quantities H_ϕ and E_θ become infinite as $kr \to 0$, but their integrals, which represent currents and voltages, would become too singular as $kr \to 0$. In the following we shall elaborate on a transmission line model to understand the implications of the above solutions.

d. Transmission Line Model

First we assume the antenna to be infinitely long and we study the outgoing TEM solution. Let $I_0(r) = Ae^{ikr}$ in (24). A voltage $V(r)$ and current $I(r)$ for the TEM mode can be defined at r as follows:

$$V_0(r) = \int_{\theta_0}^{\pi-\theta_0} d\theta\, rE_\theta = \frac{\eta}{\pi} A \left[\ln\left(\cot\frac{\theta_0}{2}\right)\right] e^{ikr}$$
$$I_0(r) = r\sin\theta_0 H_\phi(\theta = \theta_0) = Ae^{ikr}$$

The ratio of $V_0(r)$ and $I_0(r)$ gives the characteristic impedance Z_0

$$Z_0 = \frac{\eta}{\pi} \ln\left(\cot\frac{\theta_0}{2}\right) \tag{25}$$

which is a constant for all r.

Since the antenna has a finite length l, we can, for the TEM mode, model the antenna as a transmission line [Fig. 4.5.4] with length l,

characteristic impedance Z_0, and terminated with an impedance Z_t which is to be determined.

For the TEM mode on a biconical antenna with length l, we write

$$H_\phi = \frac{I_0(r)}{2\pi r \sin\theta} \tag{26a}$$

$$E_\theta = \frac{\eta}{Z_0}\frac{V_0(r)}{2\pi r \sin\theta} \tag{26b}$$

where, according to transmission line theory, the voltage along the line can be expressed in terms of the voltage $V_0(l)$ at the terminating impedance Z_t,

$$V_0(r) = V_0(l)\left[\cos k(l-r) - i\frac{Z_0}{Z_t}\sin k(l-r)\right] \tag{27a}$$

$$I_0(r) = \frac{V_0(l)}{Z_0}\left[-i\sin k(l-r) + \frac{Z_0}{Z_t}\cos k(l-r)\right] \tag{27b}$$

Thus the voltage at any point r is a superposition of the two independent solutions of sine and cosine functions of kr. Note that

$$E_\theta = \frac{1}{i\omega\epsilon r}\frac{\partial}{\partial r}(rH_\phi)$$

is satisfied and that for $r = l$, $V_0(l)/I_0(l) = Z_t$. The impedance Z_t will be shown to account for all higher-order modes inside the antenna region and the radiation field in the air region.

To include all higher-order modes, we note that along any meridian between the cones, $\nabla \times \overline{E} = 0$. We define a voltage $V(r)$

$$\begin{aligned} V(r) &= \int_{\theta_0}^{\pi-\theta_0} d\theta\, rE_\theta \\ &= V_0(r) + \frac{1}{2\pi i\omega\epsilon}\sum_u a_u \frac{d}{dr}[r\, j_u(kr)]\,[P_u(-\cos\theta_0) - P_u(\cos\theta_0)] \\ &= V_0(r) \end{aligned} \tag{28}$$

The second equality follows from (24c) and the third equality is due to the boundary condition (1). The result states that the voltage for all r along the transmission line is none other than that of the TEM mode.

The current along the line $I(r)$ is given by

$$\begin{aligned} I(r) &= 2\pi r \sin\theta_0 H_\phi(\theta = \theta_0) \\ &= I_0(r) + \sum_u a_u r j_u(kr) \sin\theta_0 \left[\frac{d}{d\theta} T_u(\theta)\right]_{\theta=\theta_0} \\ &= I_0(r) + \tilde{I}(r) \end{aligned} \tag{29}$$

where we used $\tilde{I}(r)$ to denote complementary currents due to higher-order TM modes. However, at the input end $(r = 0)$

$$I(r=0) = I_0(r=0) + \tilde{I}(r=0) = I_0(r=0) \tag{30}$$

because $r j_u(kr) \to 0$ as $r \to 0$. Thus the input current is still the same as that for the TEM modes. As a consequence, the input impedance at the antenna terminal is

$$Z_i = \frac{V(0)}{I(0)} = \frac{V_0(0)}{I_0(0)} \tag{31}$$

which depends only on the TEM mode. In terms of the terminal impedance Z_t, the input impedance Z_i is immediately determined from (27) which yields

$$Z_i = \frac{V_0(0)}{I_0(0)} = Z_0 \frac{Z_t - i Z_0 \tan kl}{Z_0 - i Z_t \tan kl} \tag{32}$$

The terminal impedance Z_t contains all the information about all the higher-order modes and the antenna configurations. To determine the terminal impedance Z_t, we now turn to the solutions in the air region and match them with those in the antenna region.

In terms of the transmission line model in Figure 4.5.4, the equivalent terminal impedance Z_t is

$$Z_t = \frac{V_0(l)}{I_0(l)} = \frac{V_0(l)}{I(l) - \tilde{I}(l)} \tag{33}$$

where $I(l)$ denotes the total current on the antenna at $r = l$ and $\tilde{I}(l)$ is the complementary current due to all modes other than TEM.

The equivalent terminal admittance Y_t of the transmission line model as shown in Figure 4.5.4 is the reciprocal of (33). For very thin antennas

$$Y_t = \frac{I(l)}{V_0(l)} - \frac{\tilde{I}(l)}{V_0(l)} \approx -\frac{\tilde{I}(l)}{V(l)} \tag{34}$$

This is because the total current at the ends of a thin antenna is vanishingly small. In fact the term $I(l)/V_0(l) = I(l)/V(l)$ is the admittance between the two caps of the antenna and as such is approximately equal to the susceptance of the electrostatic capacitance between the caps.

The complementary current $\tilde{I}(r)$ due to the higher-order TM modes ($n \neq 0$) is, from (29)

$$\tilde{I}(r) = \sum_u a_u r j_u(kr) \sin\theta_0 \left[\frac{d}{d\theta} T_u(\theta)\right]_{\theta=\theta_0}$$
$$\approx \frac{120}{Z_0} \sum_u a_u r j_u(kr) \tag{35}$$

The second equality is due to the fact that as $\theta_o \to 0$, $u \to 2m+1+\Delta$, and

$$\left[\frac{d}{d\theta} T_u(\theta)\right]_{\theta=\theta_0} \simeq -\frac{1}{2\pi} \sin u\pi \cot\frac{\theta_0}{2} \approx -\frac{\sin u\pi}{\pi\theta_0} \approx \frac{\Delta}{\theta_0} = \frac{120}{Z_0\theta_0} \tag{36}$$

The complete field solutions in the antenna and air regions are given by (24) and (19). At $r = l$, E_r should be continuous, hence

$$\sum_u u(u+1) a_u j_u(kl) T_u(\theta) = \sum_n n(n+1) b_n h_n^{(1)}(kl) P_n(\cos\theta)$$

As $\theta_0 \to 0$ and $Z_0 \to \infty$, u approaches $2m+1$ and T_u approaches $P_{2m+1}(\cos\theta)$, the limiting value of a_u is

$$\lim_{\theta_0 \to 0} a_u = \lim_{\theta_0 \to 0} a_{2m+1+\Delta} = b_{2m+1} \frac{h_{2m+1}^{(1)}(kl)}{j_{2m+1}(kl)}$$

Therefore,

$$\tilde{I}(r) = \frac{120}{Z_0} \sum_{m=0}^{\infty} b_{2m+1} \frac{h_{2m+1}^{(1)}(kl)}{j_{2m+1}(kl)} r j_{2m+1}(kr)$$

For very thin antennas, the current distribution approaches the sinusoidal distribution of the principal wave

$$I(r) = I_0 \sin k(r - l) \tag{37}$$

with

$$I_0 = i\frac{V_0(l)}{Z_0}$$

The coefficients b_{2m+1} can be obtained by comparing the expressions of the field due to the current distribution given by (37) with the field expressions in the air region. For the sinusoidal current distribution of (37) we have, in the far-field approximation,

$$H_\phi = ik\frac{e^{ikr}}{4\pi r} f_\theta$$

where

$$f_\theta = I_0 \frac{2}{k \sin\theta} [\cos(kl\cos\theta) - \cos kl]$$

The radial electric field E_r is

$$E_r = \frac{1}{-i\omega\epsilon} \frac{1}{r\sin\theta} \frac{\partial}{\partial\theta}(\sin\theta\, H_\phi) = -\eta I_0 l \frac{e^{ikr}}{2\pi r^2} \sin(kl\cos\theta) \tag{38}$$

The expansion of $\sin(kl\cos\theta)$ in terms of spherical harmonics is known to be

$$\sin(kl\cos\theta) = \sum_{m=o}^{\infty} (-1)^m (4m+3) j_{2m+1}(kl)\, P_{2m+1}(\cos\theta) \tag{39}$$

Thus (38) can be written as

$$E_r = -\eta I_o l \frac{e^{ikr}}{2\pi r^2} \sum_{m=0}^{\infty} (-1)^m (4m+3) j_{2m+1}(kl) P_{2m+1}(\cos\theta) \tag{40}$$

On the other hand, in the far-field approximation, (19*b*) becomes

$$E_r = \frac{1}{i\omega\epsilon} \frac{e^{ikr}}{2\pi r^2} \frac{1}{k} \sum_{m=0}^{\infty} 2(2m+1)(m+1) b_{2m+1} (-1)^{m+1} P_{2m+1}(\cos\theta) \tag{41}$$

due to the fact that for large arguments

$$h_n^{(1)}(kr) \approx (-i)^{n+1} \frac{e^{ikr}}{kr}$$

Equating (40) and (41) we obtain

$$b_{2m+1} = -\frac{V_0(l)}{Z_0}\frac{(4m+3)}{2(2m+1)(m+1)}k^2 l\, j_{2m+1}(kl) \tag{42}$$

and

$$\tilde{I}(r) = -\frac{60}{Z_0^2}V_0(l)\sum_{m=0}^{\infty}\frac{(4m+1)}{(2m+1)(m+1)}kl\, h_{2m+1}^{(1)}(kl)\, kr\, j_{2m+1}(kr) \tag{43}$$

therefore,

$$Y_t = -\frac{\tilde{I}(l)}{V_0(l)} = \frac{Z_a(kl)}{Z_0^2} = \frac{1}{Z_0^2}\left[R_a(kl) - iX_a(kl)\right] \tag{44}$$

where

$$R_a(kl) = 60\sum_{m=0}^{\infty}\frac{(4m+3)}{(2m+1)(m+1)}\left[kl\, j_{2m+1}(kl)\right]^2 \tag{45}$$

$$X_a(kl) = -60\sum_{m=0}^{\infty}\frac{(4m+3)}{(2m+1)(m+1)}kl\, j_{2m+1}(kl)kl\, n_{2m+1}(kl) \tag{46}$$

In terms of cosine and sine integrals

$$\begin{aligned} R_a(kl) =& 60\left[\gamma + \ln(2kl) - Ci(2kl)\right] + 30\left[\gamma + \ln(kl) - 2Ci(2kl)\right. \\ &\left. + Ci(4kl)\right]\cos(2kl) + 30\left[Si(4kl) - 2Si(2kl)\right]\sin(2kl) \end{aligned} \tag{47}$$

$$\begin{aligned} X_a(2kl) = 60Si(2kl) &+ 30\left[Ci(4kl) - \ln(kl) - \gamma\right]\sin 2kl \\ &- 30Si(4kl)\cos(2kl) \end{aligned} \tag{48}$$

where $\gamma = 0.5772\ldots$ is Euler's constant. The values of R_a and X_a are shown in Figure 4.5.5.

We note that although $Z_a = R_a - iX_a$ is not a function of the cone angle θ_0, the characteristic impedance Z_0 is. In view of (32), the input impedance Z_i of a biconical antenna is thus a function of θ_0, which is assumed to be very small in the above discussions. The characteristic impedance for the $\theta_0 = 2.7°$ antenna is 450 ohms and that for the $\theta_0 = 0.027°$ antenna is 1000 ohms. At $l/\lambda = 0.25$, the input

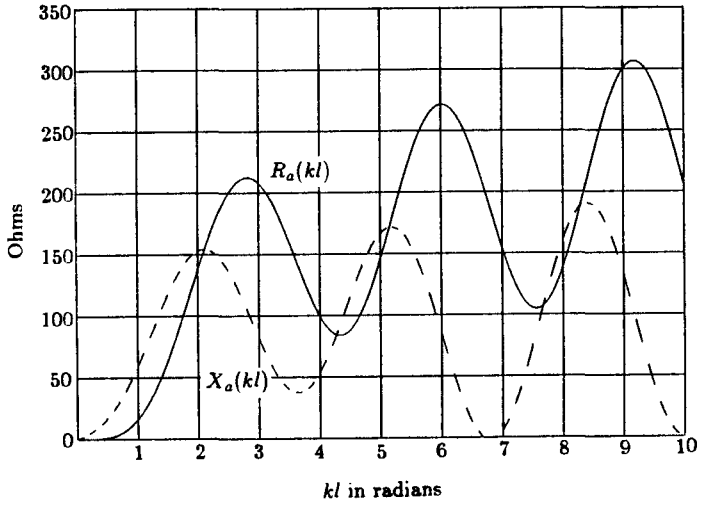


Figure 4.5.5 Resistive and reactive components of $Z_a = R_a - iX_a$.

impedance for both antennas is $Z_a(kl = \pi/2) \simeq (73.129 - i153.66)\,\Omega$. For a specific l, the input impedance of the thinner antenna as a function of frequency changes over a much wider range than that of the thicker one. This difference between thick and thin antennas is generally true and not restricted to the conical antenna. Thus, for a fixed length l, the input impedances of thick antennas are less sensitive to frequency changes and thick antennas are therefore more suitable for wide-band applications than thin antennas.

e. Formal Solution of the Biconical Antenna Problem

In the general case when θ_0 is not necessarily small, we may obtain the formal solution by making use of the orthogonal properties of the Legendre functions. The expressions for the fields in the antenna and air regions are given by (24) and (19). The problem now is to determine the coefficients a_u and b_n by matching the boundary conditions. First we note that H_ϕ in the antenna region (24*a*) may be written as

$$H_\phi(r = l) = \frac{Y_t V_0(l)}{2\pi l \sin\theta} + \frac{1}{2\pi}\sum_u a_u j_u(kl)\frac{d}{d\theta}T_u(\cos\theta) \qquad (49)$$

Multiplying both sides by $\sin\theta(dT_{u'}(\theta)/d\theta)$ and integrating from θ_0

to $\pi - \theta_0$, we obtain

$$a_u = \frac{2\pi}{N_u j_u(kl)} \int_{\theta_0}^{\pi-\theta_0} \sin\theta \, d\theta \, H_\phi(r=l) \frac{d}{d\theta} T_u(\theta) \tag{50}$$

where we have used the fact that

$$\int_{\theta_0}^{\pi-\theta_0} \sin\theta d\theta \left[\frac{d}{d\theta} T_u(\theta)\right] \left[\frac{d}{d\theta} T_{u'}(\theta)\right] = \begin{cases} 0 & \text{if } u \neq u' \\ N_u & \text{if } u = u' \end{cases}$$

However, from (19a)

$$H_\phi(r=l) = \frac{1}{2\pi} \sum_{n \text{ odd}} b_n \, h_n^{(1)}(kl) \frac{d}{d\theta} P_n(\cos\theta) \tag{51}$$

Substituting the above equation into (50) we obtain an infinite system of linear equations of the form

$$a_u = \sum_{n \text{ odd}} \alpha_{u,n} b_n \tag{52}$$

where

$$\alpha_{u,n} = \frac{h_n^{(1)}(kl)}{N_u j_u(kl)} \int_{\theta_0}^{\pi-\theta_0} \sin\theta d\theta \left[\frac{d}{d\theta} T_u(\cos\theta)\right] \left[\frac{d}{d\theta} P_n(\cos\theta)\right] \tag{53}$$

Another system of equations governing the coefficients can be obtained by considering the E_θ components at $r = l$. In the air region $E_\theta(r = l)$ is, from (19c),

$$l \, E_\theta(r=l) = \frac{1}{2\pi i\omega\epsilon} \sum_{n \text{ odd}} b_n \left\{\frac{d}{dr}\left[r h_n^{(1)}(kr)\right]\right\}_{r=l} \frac{d}{d\theta} P_n(\cos\theta) \tag{54}$$

Multiplying both sides by $\sin\theta d[P_{n'}(\cos\theta)]/d\theta$. and integrating from 0 to π we obtain

$$b_n = \frac{2n(n+1)}{2n+1} \frac{2\pi i\omega\epsilon}{\left\{d\left[r h_n^{(1)}(kr)\right]/dr\right\}_{r=l}} \int_0^\pi \sin\theta d\theta \, l E_\theta(r=l) \frac{d}{d\theta} P_n(\cos\theta) \tag{55}$$

where we have used the fact that

$$\int_0^{\pi} \sin\theta d\theta \left[\frac{d}{d\theta}P_n(\cos\theta)\right]\left[\frac{d}{d\theta}P_{n'}(\cos\theta)\right] = \begin{cases} \frac{2n+1}{2n(n+1)} & \text{if } n = n' \\ 0 & \text{if } n \neq n' \end{cases}$$

However, according to (24c)

$$lE_\theta(r=l) = \frac{\eta V_0(l)}{2\pi Z_0 \sin\theta} + \frac{1}{2\pi i\omega\epsilon}\sum_u a_u \left\{\frac{d}{dr}[rj_u(kr)]\right\}_{r=l} \frac{d}{d\theta}T_u(\theta) \tag{56}$$

for $\theta_0 < \theta < \pi - \theta_0$ and $E_\theta(r=l) = 0$ otherwise.

Therefore, substituting the above equation into (55) we obtain

$$b_n = \sum_u \beta_{n,u}a_u + K_n\frac{V_0(l)}{Z_0} \tag{57}$$

where

$$\beta_{n,u} = \frac{2n(n+1)}{2n+1}\frac{\{d[rj_u(kr)]/dr\}_{r=l}}{\left\{d[rh_n^{(1)}(kr)]/dr\right\}_{r=l}} \cdot \int_{\theta_0}^{\pi-\theta_0} d\theta \sin\theta \left[\frac{d}{d\theta}P_n(\cos\theta)\right]\left[\frac{d}{d\theta}T_u(\theta)\right] \tag{58}$$

$$K_n = i\frac{4n(n+1)}{2n+1}\frac{k}{\left\{d[rh_n^{(1)}(kr)]/dr\right\}_{r=l}}P_n(\cos\theta_0) \tag{59}$$

Therefore, the problem is formally solvable, since the coefficients a_u and b_n may be determined by the linear system of equations given by (52) and (57).

Once the coefficients are obtained, the termination admittance Y_t can be easily obtained. First, integrating (49) from θ_0 to $\pi - \theta_0$ we find

$$Y_t = \frac{\eta l}{Z_0 V_0(l)}\int_{\theta_0}^{\pi-\theta_0} d\theta\, H_\phi(r=l) \tag{60}$$

where the characteristic impedance Z_0 is given by (25). Next, substituting (51) into the above equation we obtain

$$\begin{aligned} Y_t &= \frac{-\eta}{Z_0 V_0(l)}\frac{l}{\pi}\sum_{n\ \text{odd}} b_n h_n^{(1)}(kl)P_n(\cos\theta_0) \\ &= -\frac{120}{Z_0 V_0(l)}l\sum_{n\ \text{odd}} b_n h_n^{(1)}(kl)P_n(\cos\theta_0) \end{aligned} \tag{61}$$

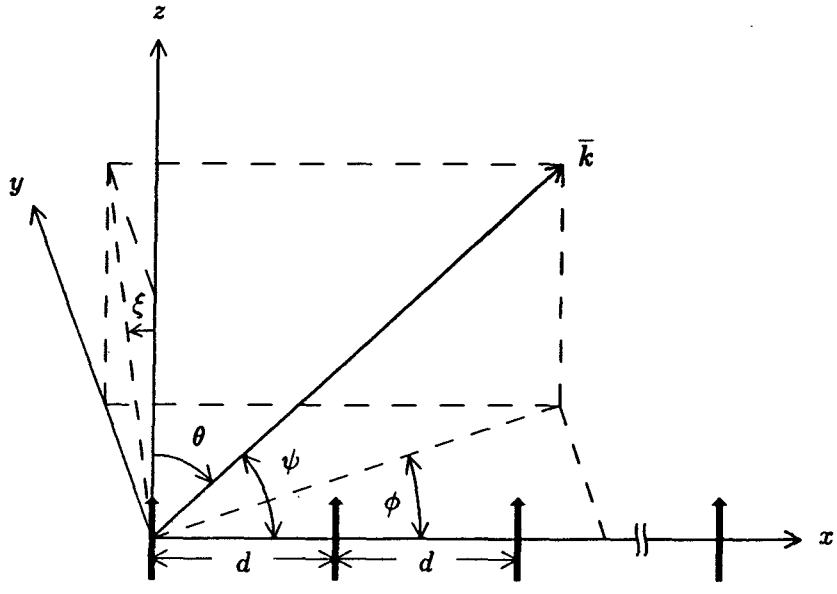


Figure 4.6.1 Linear antenna arrays.

The above equation can also be used to calculate Z_a by noting that $Z_a = Z_0^2 Y_t$.

4.6 Linear Antenna Arrays

a. Uniform Array Antenna with Progressive Phase Shift

Consider an array of N elements, pointing in the $\hat{z}$ direction and placed along the x axis with equal spacing d [Fig. 4.6.1]. Each element has a progressive phase shift α relative to its adjacent element. The current density $\overline{J}(\overline{r}')$ takes the form

$$\overline{J}(\overline{r}') = \hat{z} Il \sum_{n=0}^{N-1} e^{in\alpha} \delta(x' - nd)\delta(y')\delta(z') \tag{1}$$

The vector current moment is calculated to be

$$f(\theta, \phi) = \int dx' \int dy' \int dz' \, \overline{J}(\overline{r}') e^{in\alpha} e^{-ik(x' \sin\theta\cos\phi + y' \sin\theta \sin\phi + z'\cos\theta)}$$

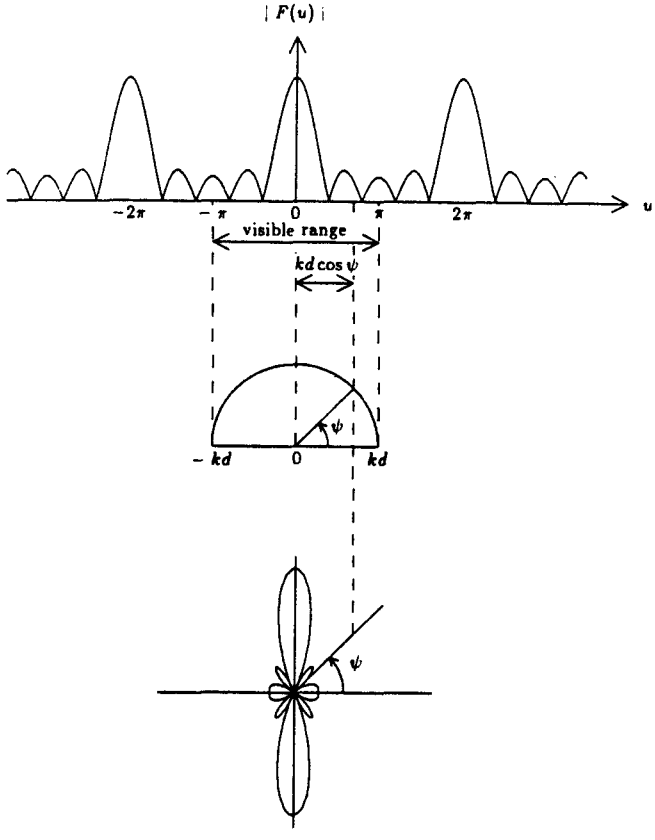


Figure 4.6.2 Array factor and radiation pattern for $N = 5$ and $kd = \pi$.

$$= \hat{z} \sum_{n=0}^{N-1} Ile^{in(kd \sin\theta \cos\phi - \alpha)} \tag{2}$$

The electric field $\overline{E}$ in the radiation zone is

$$E_\theta = i\omega\mu \frac{Ile^{ikr}}{4\pi r} \sin\theta \left[\sum_{n=0}^{N-1} e^{-in(kd\cos\psi - \alpha)} \right] \tag{3}$$

where

$$\cos\psi = \sin\theta \cos\phi \tag{4}$$

and ψ is the angle between the x axis and the position vector $\overline{r}$ [Fig. 4.6.1].

The factor in front of the square bracket in (3) is seen to be the radiation field of the individual Hertzian dipoles. The group behavior of the array is governed by the summation term known as the array factor $F(u)$,

$$F(u) = \sum_{n=0}^{N-1} e^{-inu} \tag{5}$$

where

$$u = kd\cos\psi - \alpha \tag{6}$$

The summation in the array factor is easily carried out and yields the magnitude of $F(u)$

$$|F(u)| = \left| \frac{\sin \frac{Nu}{2}}{\sin \frac{u}{2}} \right| \tag{7}$$

The magnitude of the array factor $|F(u)|$ forms a periodic pattern in u, and is plotted in Figure 4.6.2 for $N = 5$.

The part of the radiation pattern that is physically observed is determined from (6) for ψ between 0 and π which corresponds to going from the $+\hat{x}$ direction to the $-\hat{x}$ direction. This is called the visible range for which u spans the range kd and $-kd$. We can sketch the radiation pattern for the array factor as a function of the observation angle ψ.

In Figures 4.6.2 and 4.6.3 we plot the array factor as a function of u for $N = 5$. Letting $\alpha = 0$, the visible range corresponding to $d = \lambda/2$ is illustrated in Figure 4.6.2 together with the radiation pattern. As the observation angles ψ span from 0 to π, the projection $u = kd\cos\psi$ changes from kd to $-kd$ and the corresponding values for the array factor are obtained from the $|F(u)|$ plot. A similar diagram is shown in Figure 4.6.3 for $d = 4\lambda/5$. We note that these patterns are physically observed on the $x - y$ plane for which $\theta = \pi/2$. In other planes containing the x axis, the patterns must be modified by multiplication of the unit pattern of the array elements. Because the principal maximum is perpendicular to the array at $\psi = \pi/2$, such a uniform linear array is called *broadside.*

From (7) we see that the principal maxima $u_{\max}$ occur when the denominator goes to zero, namely $u_{\max}/2 = m\pi$ or

$$u_{\max} = 2m\pi \qquad m = 0, \pm 1, \pm 2, \ldots \tag{8}$$

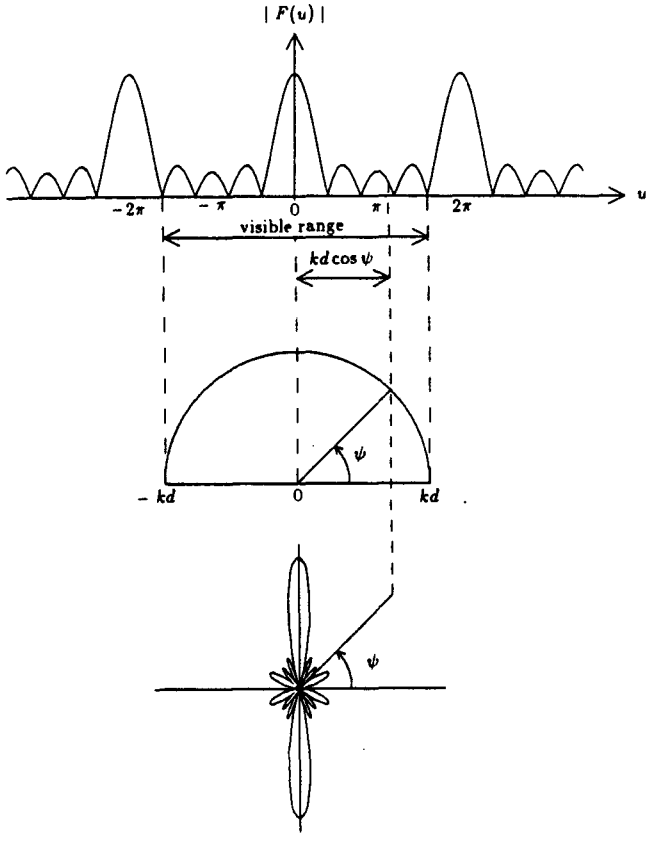


Figure 4.6.3 Array factor and radiation pattern for $N = 5$ and $kd = 8\pi/5$.

The magnitudes of these principal maxima are all equal to

$$|F(u_{\max})| = N \tag{9}$$

The location of the nulls u_n occur when the numerator goes to zero, namely $Nu_n/2 = n\pi$ or

$$u_n = \frac{2n\pi}{N} \qquad n = \pm 1, \pm 2, \ldots \tag{10}$$

which must be different from $u_{\max}$. For the five-element array shown in Figure 4.6.2, $u_n = \pm 2\pi/5$ and $\pm 4\pi/5$. The magnitude of the sidelobes can be determined from (7). First locate the positions of the maxima u_m by setting the derivative of $|F(u)|$ equal to zero and then

evaluate their magnitudes from (7). For a very large array, the sidelobe maxima can be assumed to be located midway between the nulls,

$$u_m \approx \frac{(2m+1)\pi}{N} \qquad m = 0,1,2,\ldots$$

The magnitudes are

$$|F(u_m)| = \left| \frac{\sin\left[\frac{2m+1}{2}\pi\right]}{\sin\left[\frac{2m+1}{2N}\pi\right]} \right| \approx \frac{2N}{(2m+1)\pi} \tag{11}$$

for large N within the visible range. Since the principal maxima magnitude is N, the relative magnitudes of the sidelobes are

$$\frac{|F(u_m)|}{|F(u_{\max})|} = \frac{2}{(2m+1)\pi}$$

For a large array, the first sidelobe $(m = 1)$ magnitude is $2/3\pi$ or $20\log(2/3\pi) = -13.5\,\text{dB}$ relative to the main lobe.

For broadside arrays, $\alpha = 0$ and the visible range covers $-kd \leq u \leq kd$ as ψ changes from π to 0. For a nonzero phase shift α, the visible range will be shifted to the left of the u coordinate.

Consider the case of $\alpha = kd$. We have

$$u = kd(\cos\psi - 1) \tag{12}$$

The visible range covers $-2kd \leq u \leq 0$. We illustrate the visible range and the radiation patterns in Figure 4.6.4 for $N = 5$ and $d = \lambda/2$, in Figure 4.6.5 for $N = 5$ and $d = 4\lambda/5$, and in Figure 4.6.6 for $N = 5$ and $d = \lambda/4$. We see that there may be more than one principal maximum occuring in the visible range for $kd \geq \pi/2$. A principal maximum always occurs at $\psi = 0$. Thus for $\alpha = kd$, the array is called *endfire*.

It is useful to estimate the beamwidth and directivity for both the broadside and the endfire arrays. There are two kinds of beamwidths we can define. One measures the angular spread of the main lobe in ψ between the first-null positions; the other measures the angular spread between the half-power points of the main lobe. To determine the first-null beamwidth, we find that the first null $u < 0$ occurs at

$$u_1 = kd\cos\psi_1 - \alpha = \frac{-2\pi}{N} \tag{13}$$

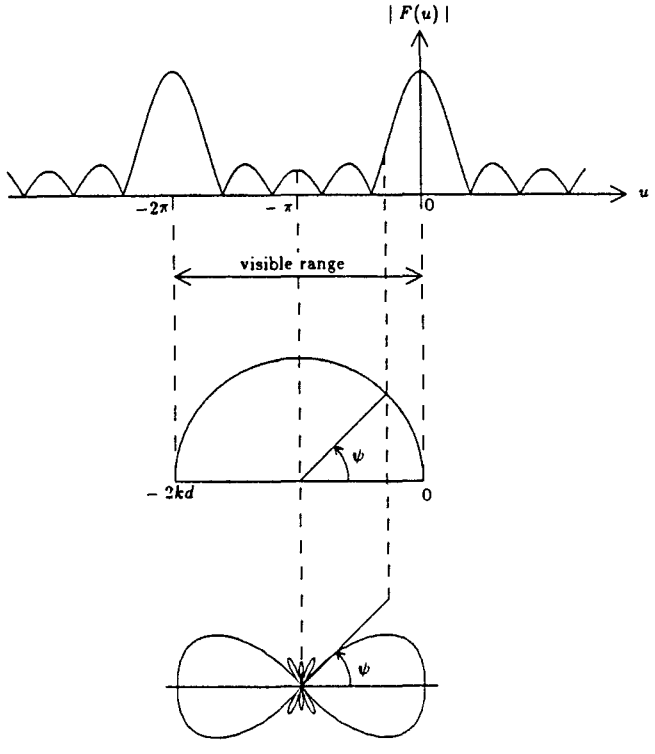


Figure 4.6.4 Visible range for $N = 5$, and $\alpha = kd = \pi$.

For a broadside array, $\alpha = 0$ and the principal maximum occurs at $\psi = \pi/2$. The beamwidth is $(\mathrm{BW})_b = 2(\psi_1 - \frac{\pi}{2})$. From (13), we see that for large N,

$$-\frac{2\pi}{Nkd} = \cos\psi_1 = -\sin\left[\psi_1 - \frac{\pi}{2}\right] \approx -\left[\psi_1 - \frac{\pi}{2}\right]$$

Thus,

$$(\mathrm{BW})_b \approx \frac{4\pi}{Nkd} \tag{14}$$

For an endfire array, $\alpha = kd$ and the principal maximum occurs at $\psi = 0$. The beamwidth is $(\mathrm{BW})_e = 2\psi_1$. From (13), we see that for large N,

$$-\frac{2\pi}{Nkd} = \cos\psi_1 - 1 \approx -\frac{\psi_1^2}{2}$$

Thus,

$$(\mathrm{BW})_e = 2\sqrt{\frac{4\pi}{Nkd}} \tag{15}$$

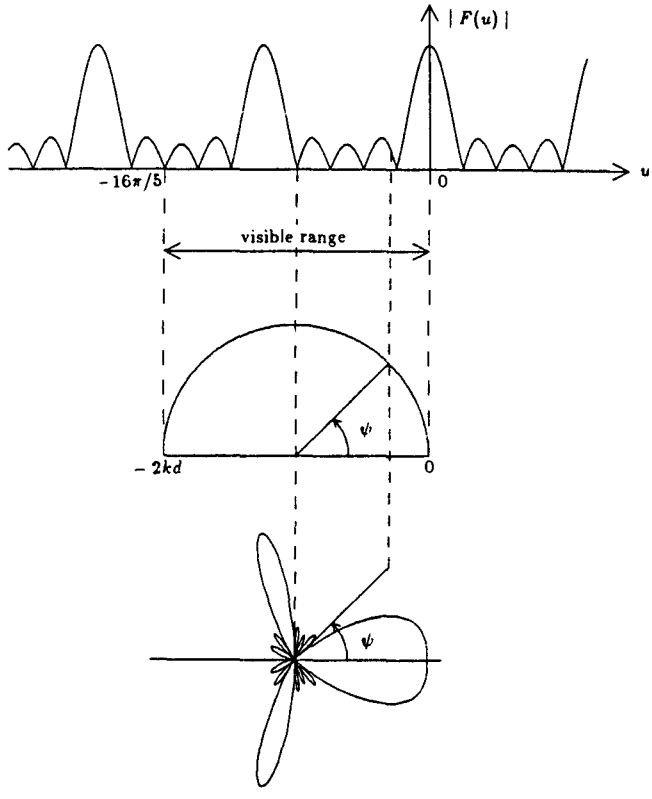


Figure 4.6.5 Visible range for $N = 5$, and $\alpha = kd = 8\pi/5$.

The first-null beamwidth for large broadside arrays is inversely proportional to N, whereas that for large endfire arrays is inversely proportional to $\sqrt{N}$. Therefore, large endfire arrays are seen to have a larger first-null beamwidth.

The directivity D is defined as the ratio of the peak power to the total radiated power distributed over 4π steradians

$$D = G(\theta,\phi)\max = \frac{4\pi\,|E(\theta,\phi)_{max}|^2}{\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta\,|E(\theta,\phi)|^2} \tag{16}$$

From (3), we have

$$E_\theta = E_0 \sin\theta \left[\sum_{n=0}^{n-1} e^{inu}\right] \tag{17}$$

where $u = kd\cos\psi - \alpha$, and $E_0 = -i\omega\mu Ile^{ikr}/4\pi r$. The directivity

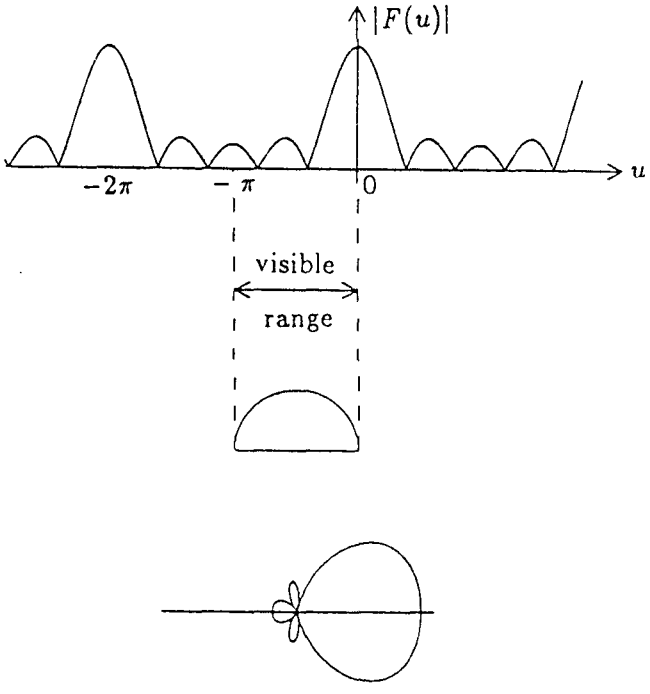


Figure 4.6.6 Visible range for $N = 5$, and $\alpha = kd = \pi/2$.

is

$$D = \frac{N^2}{\displaystyle\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin^3\theta \left[N + 2\sum_{m=1}^{N-1} (N-m)\cos mu \right]} \tag{18}$$

To facilitate the integration over 4π steradians, we use the angles ψ and ξ as shown in Figure 4.6.1 and integrate over $d\xi\, d\psi \sin\psi$ for ψ from 0 to π and ξ from 0 to 2π. The directivity is calculated as

$$D = 4\pi N^2 \left\{ \int_0^{2\pi} d\xi \int_0^{\pi} d\phi \sin\xi \,(1 - \cos^2\xi \sin^2\psi) \left[N + 2\sum_{m=1}^{N-1} (N-m)\cos mu \right] \right\}^{-1}$$

$$= 4\pi N^2 \left\{ \int_0^{\pi} d\psi \sin\psi \,(2\pi - \pi\sin^2\psi) \right.$$

$$\left[N + 2\sum_{m=1}^{N-1}(N-m)\cos mu\right]\Bigg\}^{-1}$$

$$= 4N^2\left\{\int_{-kd-\alpha}^{kd-\alpha}\frac{du}{kd}\left[1+\frac{1}{(kd)^2}(u+\alpha)^2\right]\right.$$

$$\left.\left[N + 2\sum_{m=1}^{N-1}(N-m)\cos mu\right]\right\}^{-1}$$

$$= 4N^2\left\{\frac{8}{3}N + \frac{2}{kd}\sum_{m=1}^{N-1}(N-m)\left[\frac{4}{m}\sin(mkd)\cos(m\alpha)\right.\right.$$

$$\left.\left. + \frac{4}{m^2kd}\cos(mkd)\cos(m\alpha) - \frac{4}{m^3(kd)^3}\sin(mkd)\cos(m\alpha)\right]\right\}^{-1}$$

$$= N^2\left\{\frac{2}{3}N + 2\sum_{m=1}^{N-1}(N-m)\left[\left(\frac{1}{mkd} - \frac{1}{(mkd)^3}\right)\sin(mkd)\right.\right.$$

$$\left.\left. + \frac{1}{(mkd)^2}\cos(mkd)\right]\cos(m\alpha)\right\}^{-1} \tag{19}$$

As $kd \to 0$, we have

$$D = N^2\left\{\frac{2N}{3} + \frac{2}{3}(N-m)\cos m\alpha\right\}^{-1}$$

Letting $\alpha = 0$, we obtain

$$D = N^2\left\{\frac{2N}{3} + \frac{2}{3}N(N-1)\right\}^{-1} = \frac{3}{2} \tag{20}$$

which is the directivity for a single Hertzian dipole.

b. Array Antennas with Nonuniform Current Distributions

We see that for uniform linear arrays, whether broadside or endfire, the relative power level of the first sidelobe as compared with the principal maximum is -13.5 dB and even higher for modified endfire arrays. With the use of nonuniform current excitations, the sidelobe

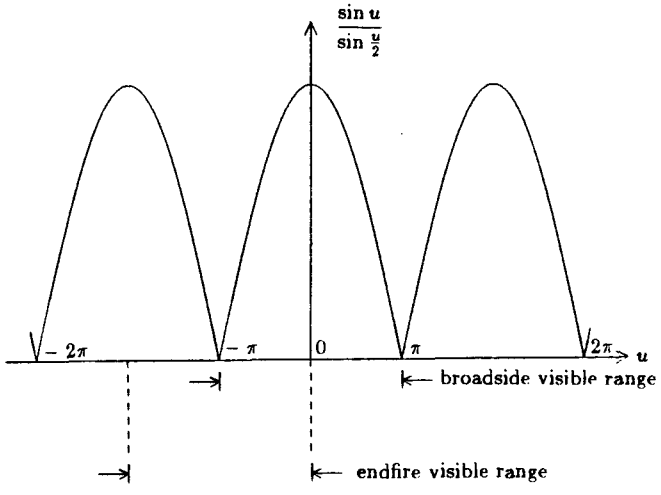


Figure 4.6.7 Array factor for $N = 2$.

levels can be reduced. One simple example is a *gabled array* which has an array factor equal to the square of that for a uniform linear array,

$$\tilde{F}(u) = \left[\sum_{n=0}^{N-1} e^{-inu}\right]^2 = 1 + 2e^{-iu} + \cdots + Ne^{-i(N-1)u} + \cdots + 2e^{-i2(N-2)u} + e^{-i2(N-1)u} \quad (21)$$

The gabled array consists of $2N-1$ elements with the center element having current amplitude N relative to the end ones having unit amplitude. The array pattern plotted against u is seen to be the square of the uniform array. The first sidelobe level is now -27 dB relative to the main lobe.

A similar construction leads to a *binomial array* which has no sidelobes. Consider a uniform linear array consisting of two elements. The array factor

$$\left|\tilde{F}(u)\right| = \left|\sum_{n=0}^{1} e^{-inu}\right| = \left|\frac{\sin u}{\sin \frac{u}{2}}\right| \quad (22)$$

If we choose d such that $kd \leq \pi$, we see from the array factor in Figure 4.6.7 (with $kd = \pi$) that neither the broadside nor the endfire

pattern has sidelobes. Taking the N th power of $\tilde{F}(u)$, we still do not have any sidelobes,

$$\left|\tilde{F}(u)\right|^N = \left|1 + \binom{N}{1} e^{-iu} + \binom{N}{2} e^{-i2u} + \cdots + \binom{N}{N} e^{-inu}\right|$$
$$= \left|\sum_{n=0}^{N} \binom{N}{n} e^{-inu}\right| \tag{23}$$

The amplitudes of the current distribution are thus equal to the coefficients of a binomial series $\binom{N}{n}$.

In general, linear arrays with nonuniform excitations can be analyzed with the z-transform method if the envelope of the current distribution is describable in functional form. As an example, consider a linear array with the sinusoidal current distribution such that

$$\bar{J}(\bar{r}') = \hat{z} I_o l \sum_{n=0}^{N-1} \sin(kax')\, \delta(x' - nd)\delta(y')\delta(z') \tag{24}$$

For symmetric excitation we also have

$$(N-1)kad = \pi \tag{25}$$

The array factor becomes

$$\tilde{F}(u) = \sum_{N=0}^{N-1} \sin(nkad) e^{-inu} \tag{26}$$

Notice that the array has only $N-2$ elements because the two end elements have zero excitation. The array factor can be cast in a standard z-transform format by letting

$$z = e^{iu} \tag{27}$$

$\tilde{F}(u)$ is then in the form of a shifted finite z-transform for the function $\sin(nkad)$. Making use of the symmetric excitation condition (25) and regarding z as a real quantity, we carry out the summation in (26) in

the following manner

$$\begin{aligned}\sum_{n=0}^{N-1}\sin(nkad)z^{-n} &= -\,Im\left\{\sum_{n=0}^{N-1}z^{-n}e^{-inkad}\right\}\\ &= -\,Im\left\{\frac{1-z^{-N}e^{-iNkad}}{1-z^{-1}e^{-ikad}}\right\}\\ &= \frac{(z^{-1}+z^{-N})\sin\left(\frac{\pi}{N-1}\right)}{1-2z^{-1}\cos\left(\frac{\pi}{N-1}\right)+z^{-2}}\end{aligned} \tag{28}$$

where we made use of the symmetry condition in (25). In fact, once the summation is cast in the z-transform format, standard formulas for many frequently encountered functions are available in tabulated forms.

The magnitude of the array factor is obtained from (26)–(28),

$$\left|\tilde{F}(u)\right| = \left|\frac{\cos\left[\frac{(N-1)u}{2}\right]\sin\left(\frac{\pi}{N-1}\right)}{\cos u-\cos\left(\frac{\pi}{N-1}\right)}\right| \tag{29}$$

The principal maximum occurs at $u=0$ for which

$$\left|\tilde{F}(u_{\max})\right| = \cot\frac{\pi}{2(N-1)} \tag{30}$$

Nulls occur at $u_n = \pm(2n+1)/(N-1)$ for $n=1,2,3,\ldots$ The value corresponding to $n=0$ or $u_0=\pi/(N-1)$ does not yield a null because both the numerator and the denominator of (29) are zero and in the limit, $\left|\tilde{F}(u_0)\right| = (N-1)/2$.

For $N=6$ and $d=\lambda/2$, the radiation pattern is shown in Figure 4.6.8. The first sidelobe maximum occurs at $u=\pm 2.44$ with $\left|\tilde{F}(2.44)\right| \Big/ \left|\tilde{F}(u_{\max})\right| = 0.12$ or -18 dB. When $d=\lambda/2$, the directivity is calculated to be $D=3.79$ for both broadside and endfire arrays.

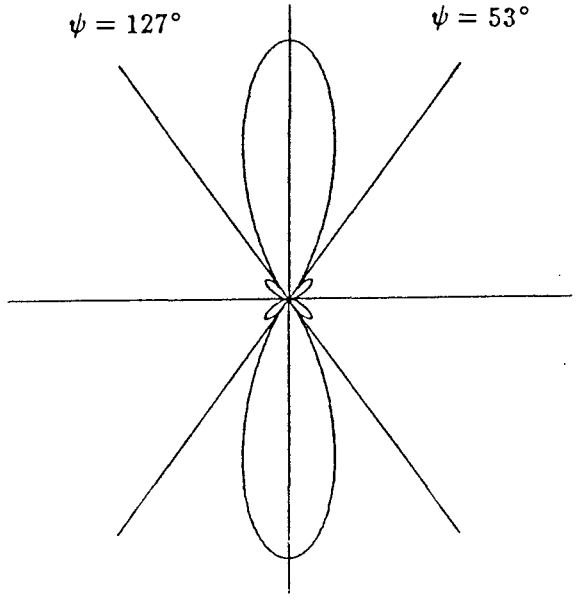


Figure 4.6.8 Radiation pattern for $N = 6$ and $d = \lambda/2$.

c. Dolph-Chebyshev Arrays

Dolph-Chebyshev arrays have equal sidelobes. The array elements are equally spaced and symmetrically excited, with a uniform progressive phase shift, and with current amplitudes determined by the coefficients of the Chebyshev polynomials. As a result of the properties of Chebyshev polynomials, such arrays provide the minimum beamwidth for a prescribed level, and conversely, the minimum sidelobe level for a prescribed beamwidth.

The Chebyshev (also spelled Tchebyscheff) polynomials are defined by

$$T_n(x) = \cos\left[n \cos^{-1} x\right] \tag{31}$$

The first few polynomials are

$$T_0(x) = 1 \tag{32a}$$
$$T_1(x) = x \tag{32b}$$
$$T_2(x) = 2x^2 - 1 \tag{32c}$$
$$T_3(x) = 4x^3 - 3x \tag{32d}$$
$$T_4(x) = 8x^4 - 8x^2 + 1 \tag{32e}$$

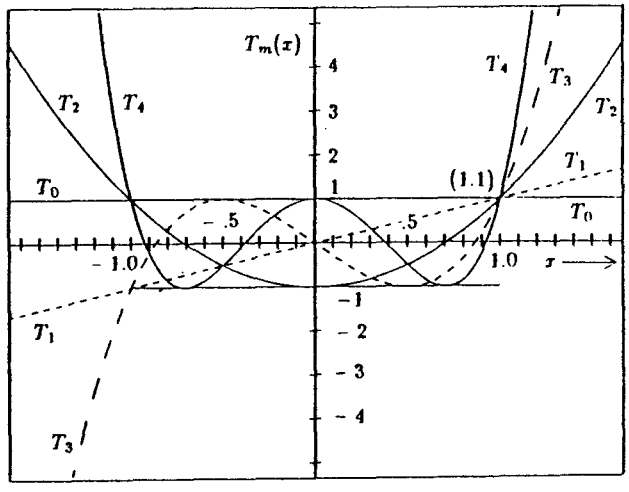


Figure 4.6.9 Chebyshev polynomials T_0, T_1, T_2, T_3 and T_4.

$$T_5(x) = 16x^5 - 20x^3 + 5x \tag{32f}$$

$$T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1 \tag{32g}$$

$$T_7(x) = 64x^7 - 112x^5 + 56x^3 - 7x \tag{32h}$$

Polynomials of higher degree can be obtained from the recurrence relation

$$T_{n+1} = 2x\,T_n(x) - T_{n-1}(x) \tag{33}$$

The polynomials $T_n(x)$ are of degree n in x. For n even, $T_n(x)$ contains only even powers of x; for n odd, $T_n(x)$ contains only odd powers of x. For $x \geq 1$, $T_n(x) = \cosh[n \cosh^{-1} x]$ and for $x < -1$, $T_n(x) = (-1)^n \cosh[n \cosh^{-1} |x|]$. Within the range $-1 \leq x \leq 1$, $T_n(x)$ oscillates between ± 1. The polynomial $T_n(x)$ passes 1 at $x = 1$ and $(-1)^n$ at $x = -1$. For $n > 0$ the zeros of $T_n(x)$ occur at

$$x_p = \cos \frac{(2p-1)\pi}{2n} \qquad p = 1, 2, \ldots, n \tag{34}$$

For $n > 1$ the extrema for $T_n(x)$ within the interval $-1 \leq x \leq 1$ occur at

$$x_m = \cos \frac{m\pi}{n} \qquad m = 1, 2, \ldots, n-1 \tag{35}$$

The Chebyshev polynomials of degrees 0 through 4 are shown in Figure 4.6.9.

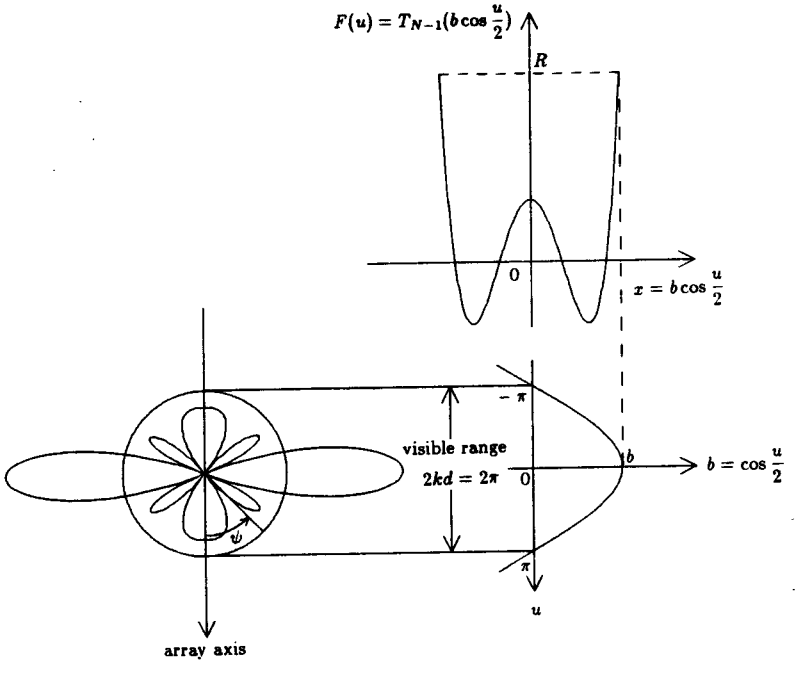


Figure 4.6.10 Broadside Dolph-Chebyshev array with $N = 5$ and $kd = \pi$.

The excitation amplitudes of an N-element Dolph-Chebyshev array are determined by equating the array factor $\tilde{F}(u)$ with the Chebyshev polynomial $T_{N-1}(x)$ under the *Dolph transformation*

$$x = b\cos\frac{u}{2} \tag{36}$$

where $b > 1$ is a parameter. We write the array factor for the Dolph-Chebyshev arrays as

$$\tilde{F}(u) = T_{N-1}\left[b\cos\frac{u}{2}\right] \tag{37}$$

where

$$u = kd\cos\psi - \alpha \tag{38}$$

As an example, we illustrate the construction of the radiation pattern for a five-element broadside Dolph-Chebyshev array with $d = \lambda/2$ and $\alpha = 0$ as shown in Figure 4.6.10. The visible range is between $u = -kd = -\pi$ and $u = kd = \pi$.

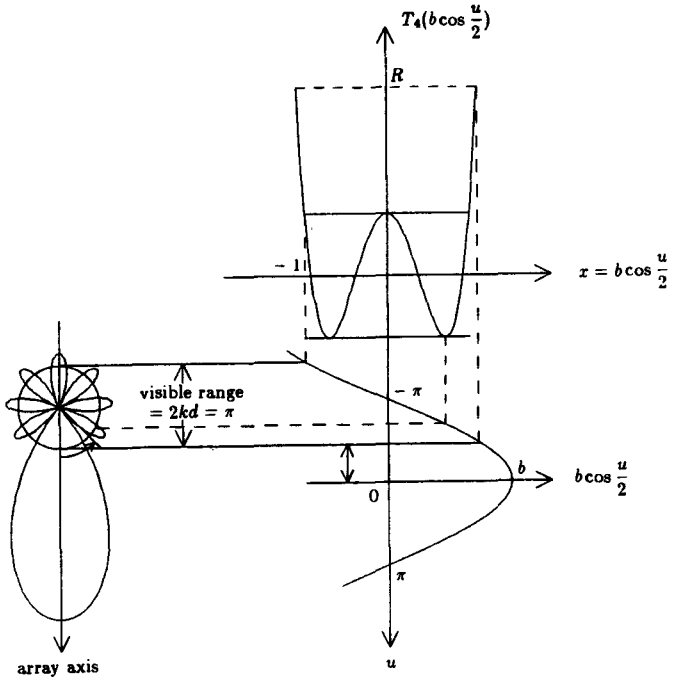


Figure 4.6.11 Dolph-Chebyshev array with $N = 5$ and $kd = \pi/2$.

Let the first-null beamwidth of the broadside array be $2[(\pi/2) - \psi_1]$. At the first null $x = b\cos[(kd\cos\psi)/2]$, we obtain from (34) by letting $p = 1$ and $n = N - 1$,

$$b\cos\frac{kd\cos\psi_1}{2} = \cos\frac{\pi}{2(N-1)} \tag{39}$$

The principal maximum occurs at $u = 0$ for which $\tilde{F}(u) = T_{N-1}(b) = R$. Notice that as $b > 1$, we have

$$R = T_{N-1}(b) = \cos[(N-1)\cos^{-1} b] = \cosh[(N-1)\cosh^{-1} b] \tag{40}$$

Thus with b specified, the beamwidth and sidelobe levels are immediately determined from (39) and (40). Conversely, if we specify the sidelobe level $1/R$, b is determined from (10)

$$b = \cosh\left[\frac{1}{N-1}\cosh^{-1} R\right]$$

and consequently, the beamwidth from (39).

As another example, we illustrate the case of a five-element endfire Dolph-Chebyshev arrays with $d = \lambda/4$ as shown in Figure 4.6.11. The principal maximum $\tilde{F}(u_{\max}) = R$ occurs at $\psi = 0$ which gives

$$u_{\max} = kd - \alpha$$

Notice that we have left the progressive phase shift α to be determined in order to ensure endfire performance. For $\psi = \pi$, we have $u = -kd - \alpha$. To include all sidelobes in the array pattern and to exclude a main lobe at $\psi = \pi$, we require that $\psi = \pi$ corresponds to $x = -1$ in $T_4(x)$. From the Dolph transformation (36) we have

$$-1 = b\cos\frac{kd + \alpha}{2} \tag{41}$$

The visible range lies between $u = kd - \alpha$ and $u = -kd - \alpha$. The construction of the radiation pattern when $d = \lambda/4$ is shown in Figure 4.6.11. Notice that the choice of $x = -1$ corresponding to $\psi = \pi$ and the choice of element separation d fix the visible range $2kd$. The variable b, which determines the amplitude of $b\cos(u/2)$ is seen to be determined by either the main lobe maximum R or the main lobe beamwidth. Main lobes at angles other than $\psi = 0$ will appear if $2kd$ becomes too large, for instance $d = \lambda/2$.

For array synthesis, suppose the sidelobe level $1/R$ is specified, we find

$$R = T_{N-1}\left[b\cos\frac{kd - \alpha}{2}\right] \tag{42}$$

Making use of (42), we obtain the equation

$$b\cos\frac{kd - \alpha}{2} = \cosh\left[\frac{1}{N-1}\cosh^{-1}R\right] \tag{43}$$

From (41) to (43), α and b can be determined.

If instead the beamwidth $2\psi_1$ is specified, we see that

$$b\cos\frac{kd\cos\psi_1 - \alpha}{2} = \cos\frac{\pi}{2(N-1)} \tag{44}$$

From (41) and (44) α and b can be determined. Sidelobe level is calculated by using (42).

We now determine the excitation amplitudes of an N-element symmetrically excited array. Placing the coordinate origin at the midpoint of the array, we write the array factor as

$$\tilde{F}(u) = \begin{cases} 2\sum_{m=1}^{N/2} a_m \cos\dfrac{(2m-1)u}{2} & N = \text{even} \qquad (45a) \\ a_0 + 2\sum_{m=1}^{(N-1)/2} a_m \cos mu & N = \text{odd} \qquad (45b) \end{cases}$$

where $u = kd\cos\psi - \alpha$ is the phase shift between adjacent array elements, and a_m are the excitation amplitudes. By the Dolph transformation, $u = 2\cos^{-1}(x/b)$. Equating the array factor to $T_{N-1}(x)$, we find

$$T_{N-1}(x) = \begin{cases} 2\sum_{m-1}^{N/2} a_m T_{2m-1}(x/b) & N = \text{even} \qquad (46a) \\ a_0 + 2\sum_{m=1}^{(N-1)/2} a_m T_{2m}(x/b) & N = \text{odd} \qquad (46b) \end{cases}$$

The excitation amplitudes a_m are determined by comparing coefficients of like powers on both sides of (46).

As an example consider a five-element Dolph-Chebyshev array. Note from the Dolph transformation (36) that, in view of (46b), we find

$$T_4(x) = a_0 + 2a_1 T_2(x/b) + 2a_2 T_4(x/b)$$

The excitation amplitudes a_0, a_1, and a_2 are determined by making use of (32c) and (32e)

$$a_0 + 2a_1\left[2(x/b)^2 - 1\right] + 2a_2\left[8(x/b)^4 - 8(x/b)^2 + 1\right] = 8x^4 - 8x^2 + 1$$

and comparing the coefficients of like powers. We find $a_2 = b^4/2$, $a_1 = 2b^4 - 2b^2$, and $a_0 = 3b^4 - 4b^2 + 1$.

To illustrate the optimum properties of the Chebyshev polynomials, let the largest value of $T_n(x)$ be at x_0, and the first zero be at x_1. For $-1 \le x \le 1$, $T_n(x)$ lies between -1 and 1. Let the visible range cover the Chebyshev polynomial from $x = -1$ to $x = x_0$. We wish to show that for a given beamwidth with first null at x_1, the sidelobe level $1/R$ is the smallest, and conversely for a given sidelobe level, the beamwidth provided by $T_n(x)$ is the smallest. To prove this

statement, we let there be another polynomial $P_n(x)$ with the same degree n and the same magnitude $R > 1$ at $x = x_0$. For the case of a given beamwidth, $T_n(x)$ and $P_n(x)$ intersect at values $x = x_0$ and $x = x_1$. If $P_n(x) \leq 1$ for $-1 \leq x \leq x_1$, then $P_n(x)$ intersects $T_n(x)$ at least $n-1$ more times. However, since specifying $n+1$ values for a polynomial of degree n uniquely determines all the coefficients of the polynomial, the two polynomials $P_n(x)$ and $T_n(x)$ must be identical. Conversely, for a given sidelobe level $1/R$, if $P_n(x)$ is to stay within the bounds ± 1 and is to possess the largest zero at x_1, such that $x_1 < x_1' < x_0$, it must intersect $T_n(x)$ at least n more times, and again $P_n(x) = T_n(x)$. This concludes the proof that $T_n(x)$ provides the optimum properties for an array pattern.

In the construction of the radiation pattern for broadside arrays, we see that if $d = \lambda/2$, the visible range will be larger than that shown in Figure 4.6.10. In fact if $d \leq \lambda$, additional main lobes will be exhibited. However, if $d < \lambda/2$, the visible range will be smaller than that shown in Figure 4.6.10, and the whole range of the Chebyshev polynomial will not be used. The Dolph-Chebyshev arrays will no longer be the optimum arrays. This can be seen for a broadside array with an odd number of elements by considering another array factor described by, instead of $T_{N-1}(x)$, another symmetric polynomial $P_{N-1}(x)$ which has the property that $P_{N-1}(b) = R$ at the principal maximum and zero at $\cos[\pi/2(N-1)]$ but with $|P_{N-1}(x)| < 1$ for $b\cos(kd/2) < x < \cos[\pi/2(N-1)]$. This is possible because $|P_{N-1}(x)|$ can be larger than unity for $0 < x < b\cos(kd/2)$, which is not in the visible range. The result is an array pattern that has the same beamwidth as a Dolph-Chebyshev array but has a smaller sidelobe level. Thus the optimum characteristic of the broadside Dolph-Chebyshev array is restricted to $d \geq \lambda/2$.

The restriction was removed by Riblet who modified the Dolph transformation (36) to $x = b\cos u + c$ for a broadside array with an odd number of elements. Notice that the Riblet transformation is not reducible to the Dolph transformation as $c = 0$. In order to determine uniquely the excitation coefficients for an $(2N+1)$ element array, the array factor is set to equal $T_n(x)$. Notice that this is in contradiction with a Dolph-Chebyshev array where the array factor for $2N+1$ elements is set to equal to $T_{2N}(x)$. The excitation coefficients

are obtained by comparing the coefficients of the equation

$$T_N(x) = a_0 + 2\sum_{m=1}^{N} a_m T_m\left[\frac{x-c}{b}\right]$$

The whole range of the Chebyshev polynomial can again be used when $d < \lambda/2$ by requiring $x = -1$ for $u = \pm kd$ or $b\cos(kd) + c = -1$. The principal maximum now occurs at $x = b + c$.

d. Array Pattern Synthesis

In array synthesis, we are asked to find a set of physically realizable excitation currents for the arrays elements for a given specification of the array pattern with regard to its beamwidth, sidelobe level, positions of nulls, directivity, and so forth. Consider the synthesis of an array having *symmetric* and nonnegative excitations with maximum possible number of nulls. The array factor takes the form

$$\tilde{F}(u) = 1 + I_1 e^{-iu} + I_2 e^{-i2u} + \ldots + I_{N-2}\, e^{-i(N-2)u} + e^{-i(N-1)u} \quad (47)$$

where we assume $I_0 = I_{N-1} = 1$ for the two end elements and $I_m = I_{N-1-m}$ for $m = 1, 2, \ldots, N-2$. The polynomial in $z = e^{-iu}$ is of the order $N-1$ roots for z. Note that $\tilde{F}(u)$ has $(N-1)/2$ unknown coefficients when N is odd and $(N-2)/2$ unknown coefficients when N is even. For odd N, we can factor $\tilde{F}(u)$ in the form

$$\tilde{F}_o(u) = \prod_{m=1}^{(N-1)/2} (1 + c_m e^{-iu} + e^{-i2u}) \quad (48)$$

which gives rise to a polynomial with symmetric coefficients. For even N we factor $\tilde{F}(u)$ in the form

$$\tilde{F}_e(u) = (1 + e^{-iu}) \prod_{m=1}^{(N-2)/2} (1 + c_m e^{-iu} + e^{-i2u}) \quad (49)$$

Note that the first factor in (49) is the result of combining the terms from symmetric positions of the polynomial. For instance, the first and last elements yield $1 + e^{-i(N-1)u}$, and, since N is an even number, this term possesses the factor $1 + e^{-iu}$.

The power patterns consist of magnitude squares of the elementary factors of $\tilde{F}(u)$. Note that, assuming real c_m,

$$(1 + c_m e^{-iu} + e^{-i2u})(1 + c_m e^{iu} + e^{i2u}) = (\xi + c_m)^2$$

where we let

$$\xi = e^{iu} + e^{-iu} = 2\cos u = 2\cos[kd\cos\psi - \alpha] \tag{50}$$

The power patterns corresponding to $\tilde{F}_o(u)$ and $\tilde{F}_e(u)$ are, in terms of the new variable ξ,

$$P_o(\xi) = \left|\tilde{F}_o\right|^2 = \prod_{m=1}^{(N-1)/2} (\xi + c_m)^2 \tag{51}$$

for odd N and

$$P_e(\xi) = \left|\tilde{F}_e\right|^2 = (\xi + 2) \prod_{m=1}^{(N-2)/2} (\xi + c_m)^2 \tag{52}$$

for even N. We see from (50) that $-2 \le \xi \le 2$; within this interval the power pattern $P(\xi)$ must be real and positive for a physically realizable array. For the maximum possible number of nulls we also have c_m real and $|c_m| < 2$ so that all zeros of the polynomial occur in the range $-2 \le \xi \le 2$.

As an example, consider the synthesis of a broadside array consisting of five elements with equal spacing of $\lambda/2$ [Ma, 1974]. The power pattern and the corresponding array factor are, respectively,

$$P_o(\xi) = (\xi + c_1)^2(\xi + c_2)^2 \tag{53}$$

$$\begin{aligned}\tilde{F}(u) &= (1 + c_1 e^{-iu} + e^{-i2u})(1 + c_2 e^{-iu} e^{-i2u}) \\ &= 1 + (c_1 + c_2)e^{-iu} + (2 + c_1 c_2)e^{-i2u} + (c_1 + c_2)e^{-i3u} + e^{-i4u}\end{aligned} \tag{54}$$

From (53) we see that nulls occur at $\xi = -c_1$ and $-c_2$. The positions of the sidelobes are determined by setting $dP_o(\xi)/d\xi = 0$ which yields $\xi_1 = -(c_1 + c_2)/2$. The values $\xi = -c_1$ and $-c_2$ are neglected because they are known to be nulls. Substituting back in (53) we find the corresponding sidelobe level

$$\frac{P[\xi_1 = -(c_1 + c_2)/2]}{P(\xi_m = 2)} = \frac{(c_1 - c_2)^4}{16(c_1 + 2)^2(c_2 + 2)^2} \tag{55}$$

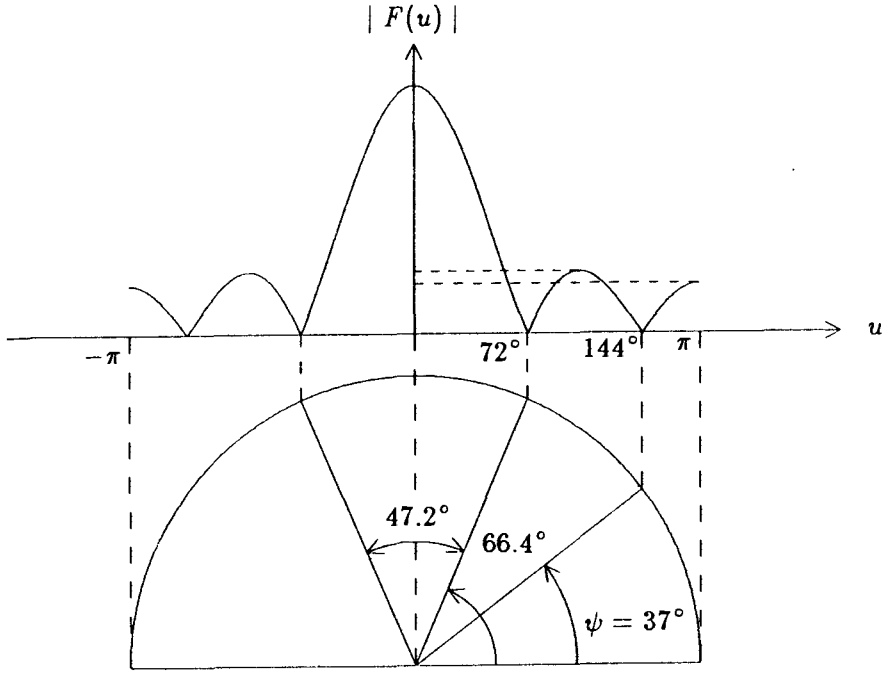


Figure 4.6.12 Array factor for Case 1.

we also regard the point $\xi = -2$ as a sidelobe maximum and find

$$\frac{P(\xi_2 = -2)}{P(\xi_m = 2)} = \frac{(c_1 - 2)^2(c_2 - 2)^2}{(c_1 + 2)^2(c_2 + 2)^2} \tag{56}$$

Case 1

We require that all excitation amplitudes be equal. From (54) we set $c_1 + c_2 = 1$ and $c_1 c_2 = -1$ and obtain $c_1 = (1 - \sqrt{5}/2)$ and $c_2 = (1 + \sqrt{5})/2$. The array factor is plotted in Figure 4.6.12. The first sidelobe level is $1/4\,(-12$ dB) and the second sidelobe level $(1/5)^2\,(-18$ dB). The first-null beamwidth is $47.2°$.

Case 2

We require all sidelobe levels be zero $(-\infty$ dB). From (55) and (56) we find $c_1 = c_2 = 2$. All sidelobe nulls are moved to $\xi = -2$ or $\xi = 0$ and π. Thus no sidelobes occur in the visible range. From (50) we see that this gives a binomial array with excitation coefficients $1 : 4 : 6 : 4 : 1$. The first-null beamwidth becomes $180°$.

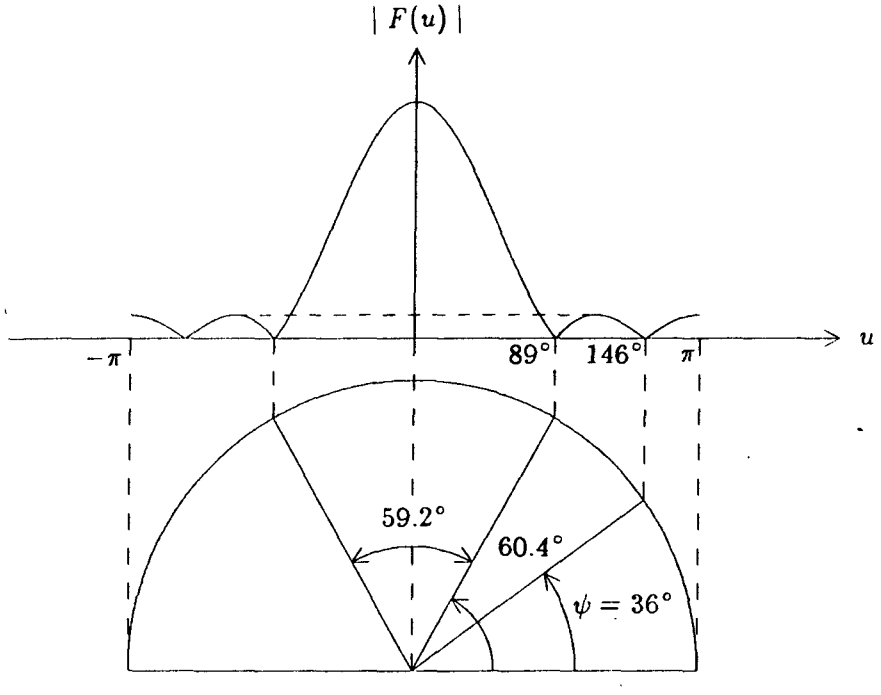


Figure 4.6.13 Array factor for Case 3.

Case 3

We require that both sidelobe levels be maintained at $-20\,\text{dB}$. From (55) and (56) we set

$$16(c_1+2)^2(c_2+2)^2 = 100(c_1-c_2)^2$$

and

$$(c_1+2)^2(c_2+2)^2 = 100(2-c_1)^2(2-c_2)^2$$

which yield $c_1 = -0.0413$ and $c_2 = 1.6498$. From (54) we find the required excitation coefficients to be $1 : 1.6085 : 1.9318 : 1.6085 : 1$. The array factor is plotted in Figure 4.6.13. The first-null beamwidth is $59.2°$. Compared with the uniform excitation, we see that the present case does give a lower sidelobe level of $(1/10)^2\,(-20\text{ dB})$ but at the expense of a larger first-null beamwidth.

When $d = \lambda/2$, the entire visible range in u is 2π. The prescribed array pattern $\tilde{F}(u)$ can be expanded into a Fourier series and approximated by a sum of the first few terms of the series. For general

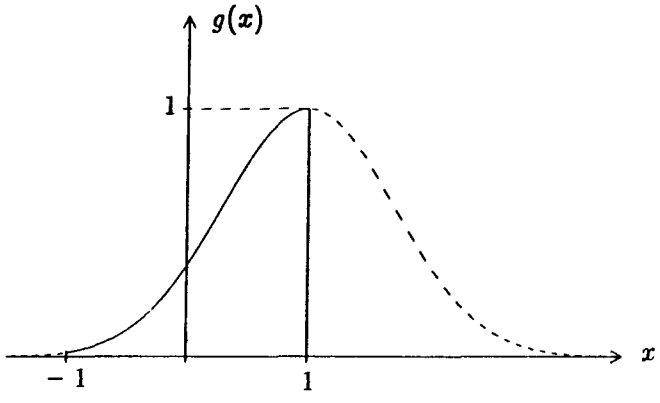


Figure 4.6.14 Plot of $g(x) = \exp[-(1-x)^2]$.

d including $d = \lambda/2$, we study, in the following, the synthesis of a linear array with a prescribed array pattern by using the method of Lagrange interpolation with Chebyshev polynomials.

A given array power pattern can be written in the form of $G(\xi)$ with $\xi = 2\cos u$ and $u = kd\cos\psi + \alpha$. As an example consider the synthesis of the array pattern [Ma, 1974]

$$G(\xi) = e^{-(\xi-2)^2/4} \qquad -2 \le \xi \le 2 \tag{57}$$

with $N = 4$

We first let $x = \xi/2$ so that

$$g(x) = e^{-(1-x)^2}$$

and the interval is normalized to $-1 \le x \le 1$ [Fig. 4.6.14]. We wish to match this function with a polynomial of degree 3. To do that, we choose the zeros of the Chebyshev polynomial $T_4(x)$ as the sampling points. We have $x_{N-1-l} = \cos[(2l+1)\pi/2\pi]$ with $l = 0, 1, 2, \ldots, N-1$ or

$$x_0 = -0.924, \quad x_1 = -0.383, \quad g(x_2) = 0.6840, \quad g(x_3) = 0.9958$$

To construct the third-degree polynomial, we apply the Lagrange interpolation formula

$$L(x) = \sum_{l=0}^{N-1} \frac{\pi(x)g(x_l)}{(x-x_l)\pi'(x_l)} \tag{58}$$

where

$$\pi(x) = (x - x_0)(x - x_1)(x - x_2)(x - x_3) \tag{59}$$

$\pi'(x)$ is the derivative of $\pi(x)$, and $L(x)$ denotes the approximating polynomial of $g(x)$ such that $L(x_i) = g(x_i)$ at $x = x_0, x_1, x_2,$ and x_3 over the normalized interval $-1 \leq x \leq 1$.

With the sampling points taken to be the zero of the Chebyshev polynomial $T_4(x)$, we find the Lagrange interpolation formula (58)

$$\begin{aligned} L(x) =& (x - x_1)(x - x_2)(x - x_3)\frac{g(x_o)}{(x_0 - x_1)(x_0 - x_2)(x_0 - x_3)} \\ &+ (x - x_0)(x - x_2)(x - x_3)\frac{g(x_1)}{(x_1 - x_0)(x_1 - x_2)(x_1 - x_3)} \\ &+ (x - x_0)(x - x_1)(x - x_3)\frac{g(x_2)}{(x_2 - x_0)(x_2 - x_1)(x_2 - x_3)} \\ &+ (x - x_0)(x - x_1)(x - x_2)\frac{g(x_3)}{(x_3 - x_0)(x_3 - x_4)(x_3 - x_2)} \\ =& -0.24628x^3 + 0.13343x^2 + 0.73586x + 0.39643 \end{aligned} \tag{60}$$

We now estimate the maximum possible error committed by the approximation of $g(x)$ with $L(x)$ as shown in (58). The accuracy of the approximation in measured by the remainder

$$R = \pi(x)\frac{g^{(N)}(x)}{N!} \qquad x_0 < x < x_{N-1}$$

where the supercript (N) denotes the Nth derivative of $g(x)$. The maximum possible error is

$$\epsilon_{max} \leq \frac{|\pi(x)|_{max}\,|g^{(N)}(x)|_{max}}{N!}$$

The maximum possible mean-square error is

$$\epsilon^2_{max} = \frac{|g^{(N)}(x)|^2_{max}}{2(N!)^2}\int_{-\infty}^{\infty} dx\,\pi^2(x)$$

Clearly the error size depends on the prescribed function $g(x)$, the number of sampling points, and their locations.

The choice of the sampling points to coincide with the zeros of $T_4(x)$ not only provides a way to determine the otherwise arbitrary sampling point positions but also makes the estimation of $|\pi(x)|_{max}$ simpler. Using the fact that $T_{N-1}(x) = 2^{N-1}(x - x_0)(x - x_1)\dots(x - x_{N-1})$ and that $|g_{N-1}(x)| \leq 1$ for $-1 \leq x \leq 1$, we find

$$\begin{aligned} |\pi(x)|_{max} &= |(x - x_0)(x - x_1)\dots(x - x_{N-1})|_{max} \\ &= \frac{1}{2^{N-1}} \left|T_{N-1}^{(x)}\right|_{max} \\ &= \frac{1}{2^{N-1}} \end{aligned}$$

We note that

$$g^{(4)}(x) = (16x^4 - 64x^3 + 48x^2 + 32x - 20)e^{-(1-x)^2}$$

Its maximum occurs at $x = 1$ giving $\left|g^{(4)}(x)\right|_{max} = 12$. We have

$$\epsilon_{max} = \frac{12}{2^3 \cdot 4!} = 0.0625$$

The actual maximum deviation between $g(x)$ and $L(x)$ occurs at $x \approx 0$ which gives an error of $\epsilon(0) = L(0) - G(0) \approx 0.0284$.

Replacing $x = \xi/2$, we find the power pattern

$$P(\xi) = -0.03078\xi^3 + 0.03331\xi^2 + 0.36787\xi + 0.39651$$

Plotted in the region $-2 \leq \xi \leq 2$, we find $P(\xi) > 0$. As $\xi \to \pm\infty$, $P(\xi) \to \mp\infty$.

For arrays with a uniformly progressive phase shift, the power pattern must satisfy the realization condition

$$P(\xi) = \sum_{m=0}^{N-1} A_m \xi^m \geq 0 \qquad \text{for } -2 \leq \xi \leq 2$$

The Weierstrass approximation theorem guarantees the existence of the coefficients A_m and the positive integer N if $g(\xi)$ is continuous for $-2 \leq \xi \leq 2$ such that for a positive quantity ϵ, $|g(\xi) - P(\xi)| < \epsilon$ for $-2 \leq \xi \leq 2$. The polynomial can be plotting for an extended

region of ξ values and is found to have a zero at $\xi = 4.4340$. We write

$$P(\xi) = -0.03078(\xi - 4.4340)(\xi^2 + 3.3350\xi + 2.9042) \tag{61}$$

There are no more zeros with real ξ for the second factor.

The power pattern must be inverted to give the array factor in order to determine the excitation coefficients. Note that by definition $\xi = 2\cos u = e^{iu} + e^{-iu}$. If the power pattern has the factor

$$P_l(\xi) = (\xi_l \pm \xi) \tag{62a}$$

then the corresponding array pattern has the factor

$$\tilde{F}_l(u) = \frac{1}{\sqrt{c_l}}(1 \pm c_l e^{-iu}) \tag{62b}$$

with $c_l = [\xi_l \pm (\xi_l^2 - 4)^{1/2}]/2$. When the power pattern has the factor

$$P_l(\xi) = (\xi + \xi_l)^2 \tag{63a}$$

with real ξ_l and $|\xi_l| < 2$, the corresponding array pattern has the factor

$$\tilde{F}(u) = 1 + \xi_l e^{-iu} + e^{-i2u} \tag{63b}$$

If the power pattern has the factor

$$P_l(\xi) = \xi^2 + 2\xi_{l1}\xi + \xi_{l1}^2 + \xi_{l2}^2 \tag{64a}$$

then the corresponding array has the factor

$$\tilde{F}_l(u) = \frac{1}{\sqrt{c_{l1}c_{l2}}}[1 + (c_{l1} + c_{l2})e^{-iu} + c_{l1}c_{l2}e^{-i2u}] \tag{64b}$$

with $c_{l2} = c_{l1}^*$ and $c_{l1} = [\xi_{l1} - i\xi_{l2} \pm ((\xi_{l1} - \xi_{l2})^2 - 4)^{1/2}]/2$ for real ξ_{l1} and ξ_{l2}. In fact we can see that for the realizable $P(\xi)$ such that $P(\xi) \geq 0$ for $-2 \leq \xi \leq 2$, the polynomial-only contains those elementary factors as listed in (61)–(63).

Returning to (54), we now find the corresponding array factor to be the multiplication of

$$\tilde{F}_1(u) = \frac{1}{2.0484}(1 - 4.1957e^{iu})$$

or

$$\frac{1}{0.4882}(1 - 0.2383e^{-iu})$$

with

$$\tilde{F}_2(u) = \frac{1}{1.3013}(1 + 2.1064e^{iu} + 1.6931e^{-i2u})$$

or

$$\frac{1}{0.7685}(1 + 1.2440e^{-iu} + 0.5906e^{-i2u})$$

Clearly, solutions for the array factor $\tilde{F}(u) = \tilde{F}_1(u)\tilde{F}_2(u)$, and therefore the excitation coefficients are not unique. With the Chebyshev interpolation method as shown above, four sets of excitation coefficients give rise to the same power pattern $P(\xi)$.

4.7 Contour Integration Methods

a. Cauchy's Theorem

The most fundamental and useful theorem in contour integration methods is Cauchy's Theorem. Consider a complex function

$$f(\alpha) = f_R(\alpha) + if_I(\alpha) \tag{1}$$

of a complex variable

$$\alpha = \alpha_R + i\alpha_I \tag{2}$$

where $f_R(\alpha)$, $f_I(\alpha)$, α_R, and α_I are all real. Assume that $f(\alpha)$ is analytic, namely its derivative exists over a domain D in the complex α plane. The boundary line of D forms a closed contour C [Fig. 4.7.1]. Cauchy's Theorem states that the line integration of $f(\alpha)$ along C is zero,

$$\oint_C d\alpha\, f(\alpha) = 0 \tag{3}$$

The direction of integration is such that when one travels along this direction, the domain D is always on his left hand side.

As an example, consider the evaluation of the integral

$$I = \int_0^\infty d\alpha \frac{\sin\alpha}{\alpha}$$

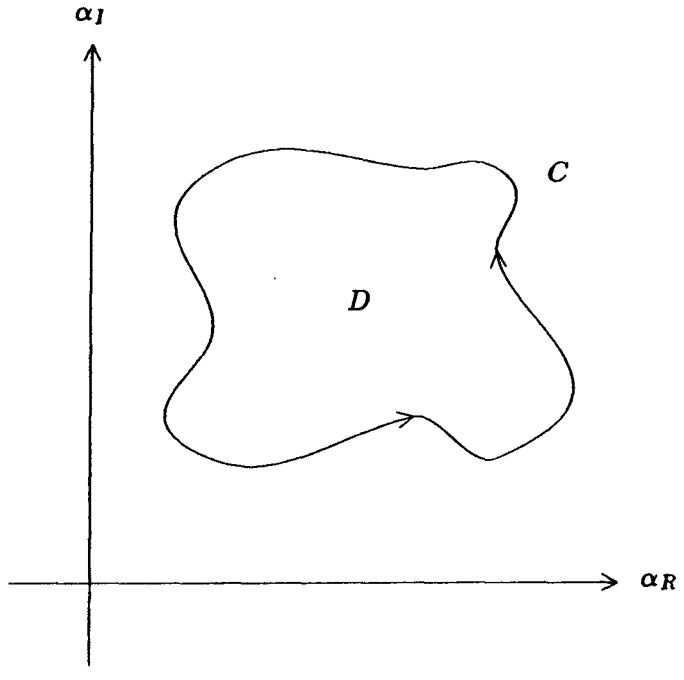


Figure 4.7.1 Contour for Cauchy's Theorem.

Writing $\sin\alpha$ in terms of exponentials, we have

$$I = \frac{1}{2i}\int_0^\infty d\alpha \left[\frac{e^{i\alpha}}{\alpha} - \frac{e^{-i\alpha}}{\alpha}\right] = \lim_{\delta\to 0}\left\{\int_\delta^\infty d\alpha \frac{e^{i\alpha}}{2i\alpha} + \int_{-\infty}^{-\delta} d\alpha \frac{e^{i\alpha}}{2i\alpha}\right\}$$

We choose a closed contour C composed of C_R, the negative real axis, C_δ, and the positive real axis as shown in Figure 4.7.2

$$I = \lim_{\substack{\delta\to 0\\ R\to\infty}} \left\{\oint_C - \int_{C_\delta} - \int_{C_R}\right\} d\alpha \frac{e^{i\alpha}}{2i\alpha}$$

where $\alpha = \alpha_R + i\alpha_I$. By Cauchy's Theorem, the first integral is zero because $e^{i\alpha}/\alpha$ is analytic inside and on contour C. The third integral also vanishes as $R \to \infty$ on account of Jordan's Lemma [Problem P4.22]. The second integral is the only one that contributes. Writing α in polar coordinates, $\alpha = \delta e^{i\phi}$, we therefore obtain

$$I = -\frac{1}{2}\lim_{\delta\to 0}\int_\pi^0 d\phi \exp(i\delta e^{i\phi}) = -\frac{1}{2}\int_\pi^0 d\phi = \frac{\pi}{2}$$

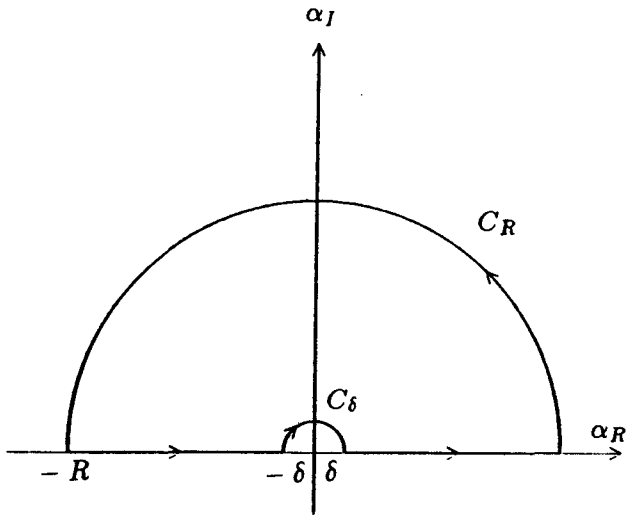


Figure 4.7.2 Contour for integration.

Notice that Cauchy's Theorem is based on the assumption that the function is analytic on the domain of integration. When there are singularities inside the domain, they must be taken into account separately. Expanding the function $f(\alpha)$ around a singularity α_0, we have

$$f(\alpha) = \sum_{n=0}^{\infty} a_n(\alpha - \alpha_0)^n + \sum_{n=1}^{\infty} \frac{a_{-n}}{(\alpha - \alpha_0)^n} \tag{6}$$

By Cauchy's Theorem no contribution comes from the first summation because it is analytical. To integrate the second summation, we first change the original contour C_0 to a new circular contour C with radius δ surrounding α_0 [Fig. 4.7.3] and let

$$\alpha - \alpha_0 = \delta e^{i\phi}$$

By Cauchy's Theorem it is seen that

$$\oint_{C_0} d\alpha f(\alpha) - \oint_C d\alpha f(\alpha) = 0$$

as the contributions from the two opposite straight lines cancel each other. Thus the integration over the original contour is equal to that

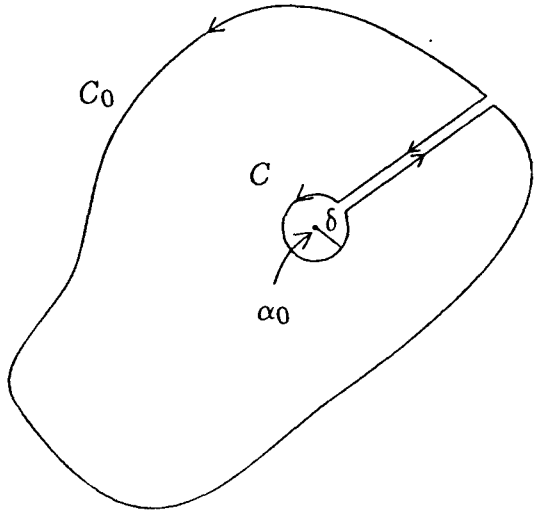


Figure 4.7.3 Contour for proof of Residue Theorem.

over C and we have

$$\oint_C d\alpha f(\alpha) = i \sum_{n=1}^{\infty} \frac{a_{-n}}{\delta^{n-1}} \int_0^{2\pi} d\phi \, e^{i(1-n)\phi}$$

Obviously, all terms other than $n = 1$ are zero. Therefore we obtain the Residue Theorem:

$$\oint_C d\alpha f(\alpha) = 2\pi i a_{-1} \tag{7}$$

Since the expansion coefficient a_{-1} is the only one left after the integration, a_{-1} is called the *residue*.

The residue of a function can be determined in many ways. If $f(\alpha)$ has a single pole, then all a_{-n} but a_{-1} vanish. We have

$$a_{-1} = \lim_{\alpha \to \alpha_o} \{(\alpha - \alpha_0) f(\alpha)\} \tag{8}$$

If $f(\alpha)$ has a pole of order m, then all subsequent terms after a_{-m} vanish, namely $a_{-l} = 0$ for $l > m$,

$$f(\alpha) = \sum_{n=0}^{\infty} a_n (\alpha - \alpha_0)^n + \frac{a_{-1}}{(\alpha - \alpha_0)} + \ldots + \frac{a_{-m}}{(\alpha - \alpha_0)^m}$$

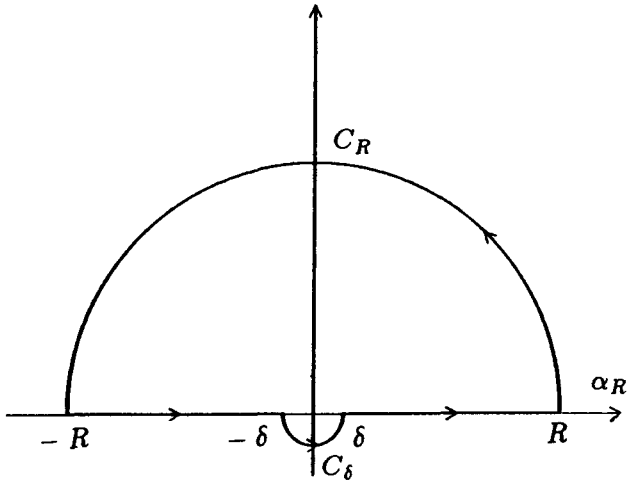


Figure 4.7.4 A closed contour including the pole at $\alpha = 0$.

and we find

$$a_{-1} = \frac{1}{(m-1)!} \lim_{\alpha \to \alpha_0} \left\{ \frac{d^{m-1}}{d\alpha^{m-1}} (\alpha - \alpha_0)^m f(\alpha) \right\} \tag{9}$$

For essential singularities, the residue can be found from known series expansions. For instance, the residue of $\exp(-1/\alpha)$ is -1.

As an example in the use of the Residue Theorem, consider again the evaluation of

$$I = \int_0^\infty d\alpha \frac{\sin \alpha}{\alpha}$$

Instead of choosing the contour as shown in Figure 4.7.2, we choose to indent around the pole at $\alpha = 0$ from below [Fig. 4.7.4]. The contour then encloses the pole which has a residue of $e^{i\alpha}/2i\alpha$ equal to $1/2i$. Thus we have

$$\begin{aligned} I &= \lim_{\substack{\delta \to 0 \\ R \to \infty}} \left\{ \oint_C - \int_{C_\delta} - \int_{C_R} \right\} d\alpha \frac{e^{i\alpha}}{2i\alpha} \\ &= 2\pi i \left[\frac{1}{2i} \right] - \lim_{\delta \to 0} \int_{C_\delta} d\alpha \frac{e^{i\alpha}}{2i\alpha} = \pi - \frac{1}{2} \int_\pi^{2\pi} d\phi = \frac{\pi}{2} \end{aligned}$$

The residue of a multi-valued function is dependent on the particular Riemann sheet where the singularity lies. To illustrate the concept of a multi-valued function and its associated Riemann sheets, consider the simple function $f(\alpha) = \sqrt{\alpha}$. The function is double-valued since, for example, as $\alpha = 1 = e^{i2m\pi}$, $f(\alpha) = e^{im\pi}$ which is $+1$ or -1 depending on whether m is even or odd. We see that at $\alpha = 0$, it is impossible to define a neighborhood in which the function is single-valued. The point $\alpha = 0$ is called a *branch point* which is another type of singularity in addition to the poles and essential singularities where the derivative of the function does not exist. To make the function single-valued, we define two Riemann sheets such that on the top Riemann sheet $4m\pi \leq \phi \leq 2(2m+1)\pi$ and on the bottom Riemann sheet $2(2m+1)\pi \leq \phi \leq 4(m+1)\pi$ where $m = 0, 1, 2, \ldots$ To separate the two Riemann sheets we choose a branch cut along the positive real α_R axis which can be visualized as creating a wedge-shaped cut [Fig. 4.7.5] such that for $0 < \phi < 2\pi$ we are on the top Riemann sheet; for $2\pi < \phi < 4\pi$, sliding to the bottom Riemann sheet, and, for $4\pi < \phi < 6\pi$, we come up again to the top sheet. It is noted that we could very well choose a different branch cut and define the two Riemann sheets differently. For instance, we may choose a branch cut along the negative α_R axis so that on the top Riemann sheet $(4m-1)\pi < \phi < (4m+1)\pi$ and on the bottom Riemann sheet $(4m+1)\pi < \phi < (4m+3)\pi$.

As an example, we find the residues for

$$f(\alpha) = \frac{\sqrt{\alpha}}{\alpha+1}$$

Notice that there is a pole at $\alpha = -1$ and a branch point at $\alpha = 0$. We choose the branch point as shown in Figure 4.7.5 except now there is a pole at $\alpha = -1$. On the top Riemann sheet, we find

$$Res. = \lim_{\alpha \to -1} \sqrt{\alpha} = e^{i(4m+1)\pi/2} = i$$

On the bottom Riemann sheet

$$Res. = \lim_{\alpha \to -1} \sqrt{\alpha} = e^{i(4m+3)\pi/2} = -i$$

If $f(\alpha)$ is analytic over a domain D bounded by contour C, then by the Residue Theorem, the function at a regular point α_0 can be

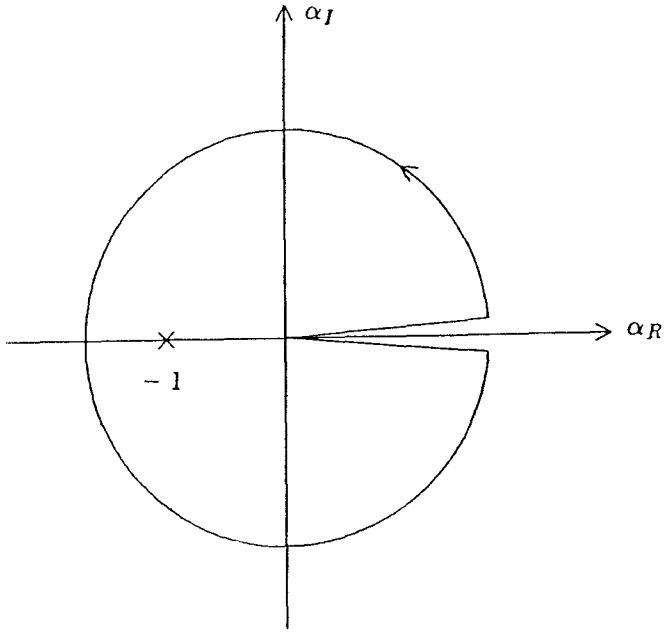


Figure 4.7.5 Contour with a branch cut.

represented by the integral

$$f(\alpha_0) = \frac{1}{2\pi i} \oint_C \frac{f(\alpha)}{(\alpha - \alpha_0)} d\alpha \tag{10}$$

which is *Cauchy's Integral Formula.* Note that we also have

$$f^{(n)}(\alpha_0) = \frac{n!}{2\pi i} \oint_C \frac{f(\alpha)}{(\alpha - \alpha_0)^{n+1}} d\alpha$$

This follows from (10) by using Leibnitz's rule of differentiation.

As another example, we derive the Kramers-Krönig relation (or the causality condition) for the real and imaginary parts of the complex permittivity for a linear, temporally dispersive medium. The linear relationship between fields $\overline{D}$ and $\overline{E}$ can be written as

$$\begin{aligned} \overline{D}(t) &= \epsilon_o \overline{E}(t) + \int_{-\infty}^{t} d\tau\, \epsilon_o \xi_e(t-\tau) \overline{E}(\tau) \\ &= \epsilon_o \overline{E}(t) + \epsilon_o \int_0^{\infty} d\tau\, \xi_e(\tau) \overline{E}(t-\tau) \end{aligned}$$

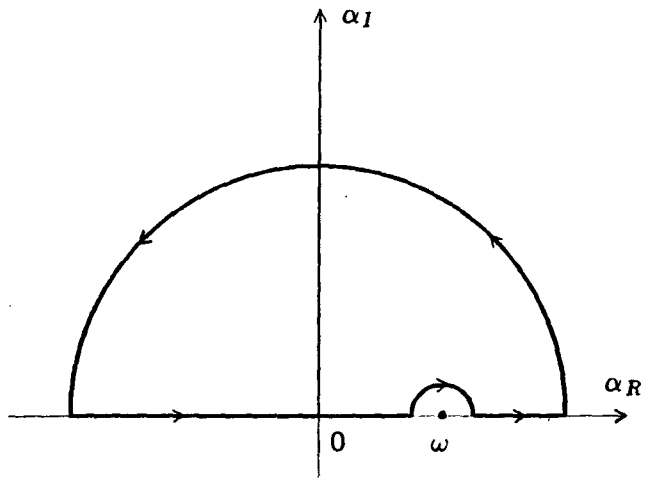


Figure 4.7.6 Contour for Kramers-Krönig's relation.

The convolution integral signifies that $\overline{D}(t)$ is determined by $\overline{E}$ at all previous times, on account of causality. Writing in terms of Fourier transformations, we have

$$\begin{aligned}\int_{-\infty}^{\infty} d\omega\, \overline{D}(\omega)e^{-i\omega t} &= \epsilon_o \int_{-\infty}^{\infty} d\omega\, \overline{E}(\omega)e^{-i\omega t} \\ &\quad + \epsilon_o \int_0^{\infty} d\tau\, \xi_e(\tau) \int_{-\infty}^{\infty} d\omega\, \overline{E}(\omega)e^{-i\omega(t-\tau)} \\ &= \epsilon_o \int_{-\infty}^{\infty} d\omega \left[1 + \int_0^{\infty} d\tau\, \xi_e(\tau)e^{i\omega\tau}\right] \overline{E}(\omega)e^{-i\omega t}\end{aligned}$$

with the time dependence $e^{-i\omega t}$ for the fields, the permittivity $\epsilon(\omega)$ is seen to be

$$\epsilon(\omega) = \epsilon_o \left[1 + \int_0^{\infty} d\tau\, \xi_e(\tau)e^{i\omega\tau}\right]$$

Notice that because of causality, the integration range for τ is from 0 to ∞. The susceptibility $\xi_e(\tau)$ is assumed to be finite throughout the range of integration, and $\epsilon(\omega)$ is a single-valued regular function over the upper half of the complex ω plane. From the defining equation for $\epsilon(\omega)$, we see that $\epsilon^*(\omega) = \epsilon(-\omega)$.

Using Cauchy's Theorem we integrate $\oint_C d\alpha\, [\epsilon(\alpha) - \epsilon_\infty]/(\alpha - \omega)$ over a closed contour consisting of a semicircle of infinite radius with

the straight side along the real axis but indented around the point $\alpha = \omega$ [Fig. 4.7.6]. The integral over the semicircle of infinite radius vanishes in view of Jordan's Lemma. Defining the principal value PV of the integral to be the result of the integration along the real axis except at $\alpha = \omega$, we find

$$\mathrm{PV}\int_{-\infty}^{\infty} d\alpha \frac{\epsilon(\alpha) - \epsilon_\infty}{\alpha - \omega} - i\pi[\epsilon(\omega) - \epsilon_\infty] = 0$$

Separating the real and imaginary parts, we find the Kramers-Krönig's relationship or the causality condition.

$$\epsilon_R(\omega) - \epsilon_\infty = \frac{1}{\pi}\mathrm{PV}\int_{-\infty}^{\infty} d\alpha \frac{\epsilon_I(\alpha)}{\alpha - \omega} \tag{11a}$$

$$\epsilon_I(\omega) = -\frac{1}{\pi}\mathrm{PV}\int_{-\infty}^{\infty} d\alpha \frac{\epsilon_R(\alpha) - \epsilon_\infty}{\alpha - \omega} \tag{11b}$$

Equation (11*a*) is known mathematically as the Hilbert transform relation that relates the real part of $\epsilon(\omega) - \epsilon_\infty$ to its imaginary part. Equation (11*b*) is the inverse Hilbert transform relation.

b. Asymptotic Series for Hankel Functions

The saddle point method is a useful tool in evaluating asymptotic values for an integral when its integrand contains a very large parameter. In the following we illustrate the method by finding the asymptotic values for the Hankel functions. Using wave concepts we argue first that a cylindrical wave can be represented by a superposition of plane waves emerging from all angles, real and complex. Then we define Hankel functions in terms of a contour integration on a complex plane [Sommerfeld, 1962].

The wave equation in cylindrical coordinates takes the form

$$\left[\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2} + k^2\right] u(\rho,\phi) = 0 \tag{12}$$

The solution to this equation is

$$u(\rho,\phi) = H_n(k\rho)e^{\pm in\phi}$$

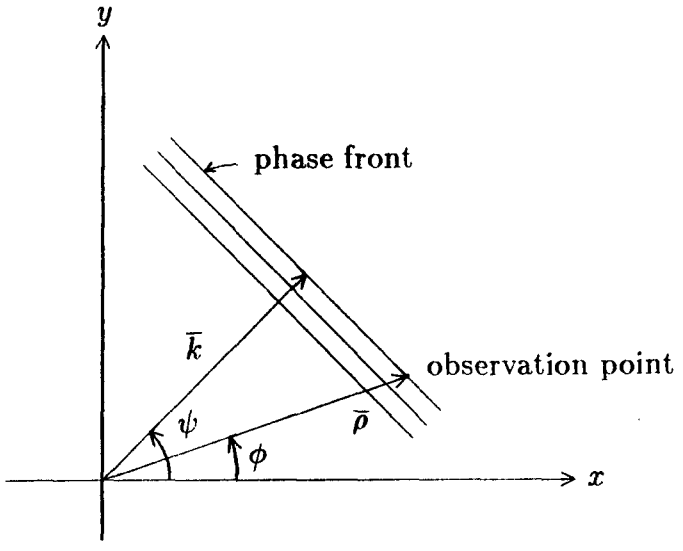


Figure 4.7.7 Cylindrical wave as superposition of plane waves.

where $H_n(k\rho)$ is the Hankel function. The same wave equation in rectangular coordinates takes the form

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + k^2\right] \tilde{u}(x, y) = 0$$

The two wave equations are related by a coordinate transformation. The familiar plane wave solution to the above equation is

$$\tilde{u}(x, y) = e^{ik_x x + ik_y y}$$

In cylindrical coordinates, the plane wave becomes

$$e^{ik_x x + ik_y y} = e^{ik\rho\cos(\psi - \phi)}$$

where

$$\bar{k} = \hat{x}k_x + \hat{y}k_y = \hat{x}k\cos\psi + \hat{y}k\sin\psi$$
$$\bar{\rho} = \hat{x}x + \hat{y}y = \hat{x}\rho\cos\phi + \hat{y}\rho\sin\phi$$

The wave vector $\bar{k}$ indicates the direction of the plane wave propagation, and the position vector $\bar{\rho}$ represents the observation point [Fig. 4.7.7].

We can view a cylindrical wave as a superposition of plane waves, uniform and nonuniform, emerging from all angles ψ, real and complex. Denoting wave amplitudes by $C_n e^{in\psi}$, we write

$$u(\rho,\phi) = \int_\Gamma d\psi\, C_n\, e^{in\psi} e^{ik\rho\cos(\psi-\phi)} \tag{13}$$

It is straightforward to show that this integral is indeed a solution to the wave equation (12) in cylindrical coordinates. The path of integration Γ and the amplitude constant C_n are still to be specified.

The path Γ on the complex ψ plane must be chosen to assure that this integral converges properly. To put the solution into the desired form, we set $\alpha = \psi - \phi$. Thus

$$u(\rho,\phi) = e^{in\phi} \int_\Gamma d\alpha\, C_n e^{ik\rho\cos\alpha + in\alpha} \tag{14}$$

We now investigate the convergence of the integral in the neighborhood of infinity. On the complex plane $\alpha = \alpha_R + i\alpha_I$, the term in the exponent takes the form

$$ik\rho\cos\alpha + in\alpha = i\,[k\rho\cos\alpha_R\cosh\alpha_I + n\alpha_R] + [k\rho\sin\alpha_R\sinh\alpha_I - n\alpha_I] \tag{15}$$

As $\alpha_I \to +\infty$, the real part is positive for $\sin\alpha_R > 0$ and negative for $\sin\alpha_R < 0$. As $\alpha_I \to -\infty$, the real part is positive for $\sin\alpha_R < 0$ and negative for $\sin\alpha_R > 0$. When the real part is positive, the integrand diverges exponentially as $\alpha_I \to \pm\infty$, and we denote with shaded regions in Figure 4.7.8. We let $\xi = k\rho$. The Hankel function of the first kind is defined as

$$H_\nu^{(1)}(\xi) = \frac{1}{\pi}\int_{\Gamma_1} d\alpha\, e^{i(\xi\cos\alpha + \nu\alpha - \nu\pi/2)} \tag{16}$$

where the integration path Γ_1 is shown in Figure 4.7.8. We see that the real part in (15) goes to $-\infty$ for $\alpha_R = -\pi/2$ and $\alpha_I \to +\infty$, and for $\alpha_R = \pi/2$ and $\alpha_I \to -\infty$. Thus the integral converges. The Hankel function of the second kind is defined by the same integral but follows a different path Γ_2,

$$H_n^{(2)}(\xi) = \frac{1}{\pi}\int_{\Gamma_2} d\alpha\, e^{i\xi(\cos\alpha + n\alpha - n\pi/2)} \tag{17}$$

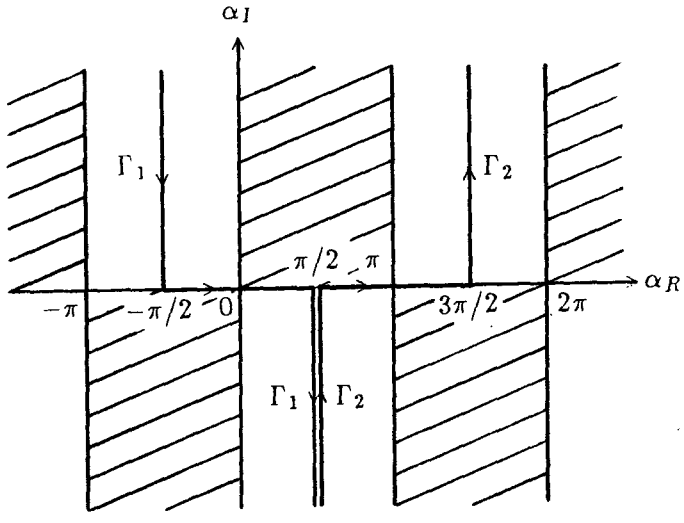


Figure 4.7.8 Integration paths for Hankel functions.

The path of integration Γ_2 is also shown in Figure 4.7.8.

We now determine asymptotic values of $H_n^{(1)}(\xi)$ as $\xi \to \infty$ by using the saddle point method. In the limit of $\xi \to \infty$, the integrand is dominated by the exponential term $e^{i\xi \cos\alpha}$. Saddle points are found from

$$\frac{d}{d\alpha}(\xi \cos\alpha) = 0$$

which gives $\alpha = n\pi$, $n = 0, \pm 1, \pm 2, \ldots$ Thus for the Hankel function of the first kind $H_n^{(1)}(\xi)$ for which $-\pi/2 \leq \alpha \leq \pi/2$, a saddle point occurs at $\alpha = 0$. For the Hankel function of the second kind, the saddle point will be at $\alpha = \pi$. We note that away from the saddle point the integrand ascends into the shaded regions and descends into the unshaded regions. Since integration may be viewed as finding the area under a curve, we can search for a new integration path passing the saddle point in such a way that most of the contribution to the integrand comes from a small portion of the new path near the saddle point. By Cauchy's Theorem we can deform the original path of integration Γ_1 to the new integration path.

The second step in the saddle point method is to determine the new integration path by requiring that on it the imaginary part of the exponent $i\xi \cos\alpha = i\xi \cos\alpha_R \cosh\alpha_I + \xi \sin\alpha_R \sinh\alpha_I$ be a constant

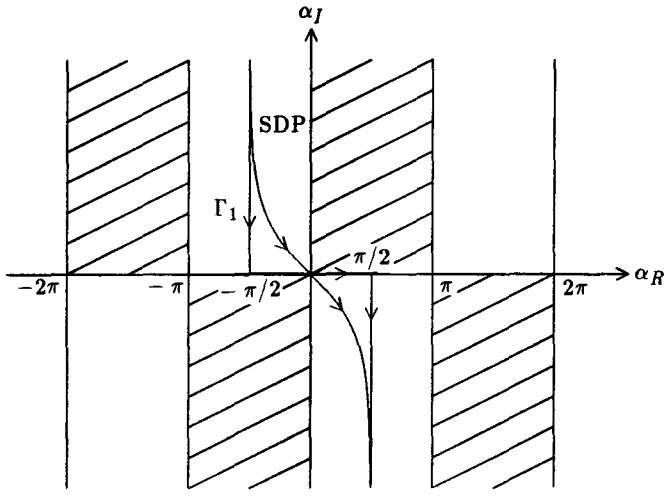


Figure 4.7.9 Integration paths on α plane.

equal to its value at the saddle point. We find

$$\cos \alpha_R \cosh \alpha_I = 1 \tag{18}$$

This choice eliminates the oscillatory behavior caused by the imaginary part in the exponential term $e^{i\xi \cos \alpha}$. The real part of the exponential term has its maximum value at the saddle point and decreases most rapidly along this path. Thus the major contribution to the integral comes from near the saddle point. The new integration path is called the steepest descent path (SDP). We see from (18) that as $\alpha_I \to \pm\infty$, $\alpha_R \to \mp\pi/2$ so that the ends of SDP meet Γ_1 as shown in Fig. 4.7.9. Near the saddle point, we expand (18) around $\alpha = 0$ to obtain

$$\left(1 - \frac{1}{2}\alpha_R^2 + \ldots\right)\left(1 + \frac{1}{2}\alpha_I^2 + \ldots\right) = 1$$

which gives $\alpha_R \approx \pm\alpha_I$. We choose the negative sign which gives $\alpha_R = -\alpha_I$. Thus the steepest descent path passes through the saddle point at $-\pi/4$ with respect to the α_R axis [Fig. 4.7.9].

The third step is to deform the original path of integration to the steepest descent path. For the Hankel function of the first kind

$H_\nu^{(1)}(\xi)$, no singularity is crossed in the process of deformation. To evaluate the saddle-point contributions of the integral, we let

$$-s^2 = i(\cos\alpha - 1) \tag{19}$$

such that $s = 0$ at $\alpha = 0$. Along the SDP, $-s^2 = \sin\alpha_R \sinh\alpha_I$ is always negative real. The integral (11) becomes

$$H_\nu^{(1)}(\xi) = \frac{1}{\pi} e^{i(\xi - \nu\pi/2)} \int_{-\infty}^{\infty} ds \frac{d\alpha}{ds} e^{i\nu\alpha} e^{-\xi s^2}$$

Since contributions to the integral come from near the saddle point, we may expand (19) around $\alpha = 0$ to obtain $s^2 = i\alpha^2/2$ which yields

$$\frac{d\alpha}{ds} = \sqrt{2}\, e^{-i\pi/4}$$

where s is a real variable. The integral becomes, near $\alpha \approx 0$,

$$H_\nu^{(1)}(\xi) \approx \frac{\sqrt{2}}{\pi} e^{i(\xi - \nu\pi/2)} \int_{-\delta}^{\delta} ds\, e^{-i\pi/4} e^{-\xi s^2}$$

where δ is a small number. Since away from $\alpha = 0$, $e^{-\xi s^2}$ decays very rapidly, we may replace δ by ∞ and make use of the formula

$$\int_{-\infty}^{\infty} ds\, e^{-\xi s^2} = \sqrt{\frac{\pi}{\xi}}$$

We thus determine the asymptotic value for $H_\nu^{(1)}(\xi)$ to be

$$H_\nu^{(1)}(\xi) \approx \sqrt{\frac{2}{\pi\xi}} e^{i(\xi - \nu\pi/2 - \pi/4)} \tag{20}$$

This result is to the order of $\xi^{-1/2}$. The saddle-point method can be further used to obtain an asymptotic series for $H_\nu^{(1)}(\xi)$ in inverse powers of ξ and for the case of ν comparable to or larger than ξ.

d. Saddle-Point Method

The saddle-point method as outlined in the above section for the computation of asymptotic values for $H_\nu^{(1)}(\xi)$ can be generalized to find asymptotic series for integrals of the form

$$I(\xi) = \int_\Gamma d\alpha\, F(\alpha) e^{\xi f(\alpha)} \tag{21}$$

where ξ is a large parameter. Let the saddle point be at $\alpha = \alpha_0$ which is determined from $f'(\alpha) = 0$ where we use a prime to denote the derivative of the function. Let Γ be the steepest descent path determined by $f_I(\alpha) = f_I(\alpha_0)$. The contribution to the saddle point is computed with the transformation

$$-s^2 = f(\alpha) - f(\alpha_0) \tag{22}$$

On the s plane, the integral becomes

$$I(\xi) = e^{\xi f(\alpha_0)} \int_{-\infty}^{\infty} ds\, \Phi(s) e^{-\xi s^2} \tag{23}$$

where

$$\Phi(s) = \frac{d\alpha}{ds} F[\alpha(s)] \tag{24}$$

The idea is to first expand α near the saddle point α_0 in a Taylor series,

$$\alpha(s) - \alpha_0 = \sum_{n=1}^{\infty} a_n s^n \tag{25}$$

then expand

$$\Phi(s) = \sum_{m=0}^{\infty} A_m s^m \tag{26}$$

and make use of the formulas

$$\int_{-\infty}^{\infty} ds\, s^{2m+1} e^{-\xi s^2} = 0 \tag{27a}$$

$$\int_{-\infty}^{\infty} ds\, s^{2m} e^{-\xi s^2} = \frac{(2m)!}{m! 2^{2m} \xi^m} \sqrt{\frac{\pi}{\xi}} \tag{27b}$$

to obtain the asymptotic series for $I(\xi)$ in inverse powers of $\xi^{-1/2}$.

We now illustrate the solution to orders of $\xi^{-3/2}$. The coefficients a_n in (25) are determined from (22) by expanding $f(\alpha)$ about $f(\alpha_0)$

$$\begin{aligned}
-s^2 &= \frac{1}{2!}f''(\alpha_0)(\alpha-\alpha_0)^2 + \frac{1}{3!}f'''(\alpha_0)(\alpha-\alpha_0)^3 \\
&\qquad + \frac{1}{4!}f^{iv}(\alpha_0)(\alpha-\alpha_0)^4 + \cdots \\
&= \frac{1}{2}f''(\alpha_0)\left[a_1^2 s^2 + 2a_1a_2 s^3 + (a_2^2 + 2a_1a_3)s^4 + \cdots\right] \\
&\qquad + \frac{1}{6}f'''(\alpha_0)\left[a_1^3 s^3 + 3a_1^2 a_2 s^4 + \cdots\right] \\
&\qquad + \frac{1}{24}f^{iv}(\alpha_0)\left[a_1^4 s^4 + \cdots\right] + \cdots
\end{aligned}$$

Comparison of coefficients yields

$$a_1 = \sqrt{\frac{-2}{f''}} \tag{28a}$$

$$a_2 = -\frac{f'''}{6f''}a_1^2 \tag{28b}$$

$$a_3 = \frac{1}{24}\left[\frac{5}{3}\left(\frac{f'''}{f''}\right)^2 - \frac{f^{iv}}{f''}\right]a_1^3 \tag{28c}$$

$$\vdots \qquad\qquad \vdots$$

The coefficients A_n in (26) are determined from (24) which gives

$$\begin{aligned}
\Phi(s) &= \frac{d\alpha}{ds}\sum_{k=0}^{\infty}\frac{1}{k!}F^{(k)}(\alpha_0)(\alpha-\alpha_0)^k \\
&= \sum_{k=0}^{\infty}\frac{1}{k!}F^{(k)}(\alpha_0)\left[\sum_{n=1}^{\infty}a_n s^n\right]^k \cdot \left[\sum_{m=1}^{\infty}m a_m s^{m-1}\right] \\
&= F(\alpha_0)(a_1 + 2a_2 s + 3a_3 s^2 + \cdots) \\
&\qquad + F'(\alpha_0)(a_1 s + a_2 s^2 + \cdots)(a_1 + 2a_2 s + \cdots) \\
&\qquad + \frac{1}{2}F''(\alpha_0)(a_1^2 s^2 + \cdots)(a_1 + \cdots) + \cdots
\end{aligned}$$

Comparison of coefficients yields

$$A_0 = a_1 F(\alpha_0) \tag{29a}$$

$$A_2 = 3a_3F(\alpha_0) + 3a_1a_2F'(\alpha_0) + \frac{1}{2}a_1^3F''(\alpha_0) \tag{29b}$$

$$\vdots$$

We do not care about the coefficients $A_1, A_3, \ldots$ because by virtue of (27a), they disappear in the final solutions.

Substituting values of a_n from (28) in (29) and making use of the formulas (27), the integral in (23) gives

$$I(\xi) = F(\alpha_0)e^{\xi f(\alpha_0)}\sqrt{\frac{2\pi}{-\xi f''}}\left\{1 + \frac{1}{2\xi f''}\left[\frac{f'''}{f''}\frac{F'}{F} + \frac{1}{4}\frac{f^{iv}}{f''} - \frac{5}{12}\frac{(f''')^2}{(f'')^2} - \frac{F''}{F}\right] + \cdots\right\} \tag{30}$$

Note that the procedure illustrated can be used to determine the coefficients a_n and A_n up to any order and that $I(\xi)$ can be expanded in higher inverse powers of $\xi^{-1/2}$. This series, which is asymptotic, diverges for any fixed ξ. In spite of its divergent behavior, an asymptotic series is very useful. The sum of the first few terms of the series approaches the values of the function that the series represents, and then diverges as more terms are added. The error introduced in representing the function by the first n terms is of the order of the $(n+1)$ th term. When the first few terms of the asymptotic series converge to the actual value of the function, the convergence is much faster than a convergent series expansion of the function. The first term of the asymptotic series may be considered as the leading behavior of the integral. The leading behavior of an integral can be evaluated readily by the saddle-point method. When the integral is expressed along either the real or the imaginary axis and can be evaluated without deforming the path of integration, the methods of Laplace and of stationary phase are also very convenient.

As an example, we find the saddle-point contribution to the integral

$$I(\xi) = \int_\Gamma d\alpha\, F(\alpha)e^{i\xi\cos(\alpha-\alpha_0)} \tag{31}$$

as $\xi \to \infty$. We let

$$f(\alpha) = i\cos(\alpha - \alpha_0)$$

and the saddle point occurs at $\alpha = \alpha_0$. Making use of (30), we immediately obtain

$$I(\xi) = F(\alpha_0)e^{i\xi}\sqrt{\frac{2\pi}{i\xi}}\left\{1 - \frac{i}{2\xi}\left[\frac{1}{4} + \frac{F''}{F}\right] + \cdots\right\} \tag{32}$$

It is noted that, if the integration path Γ is not the steepest descent path, then in deforming from Γ to the SDP which is determined from $Im[f(\alpha)] = Im[f(\alpha_0)]$, there may be singularities lying between the new and the old paths of integration. Contributions attributed to such singularities must be taken into account separately.

As another example we find the leading behavior of the Hankel function $H_\nu^{(1)}(\xi)$ when both ξ and ν are large,

$$H_\nu^{(1)}(\xi) = \frac{1}{\pi}\int_{\Gamma_1} d\alpha\, e^{i(\xi\cos\alpha + \nu\alpha - \nu\pi/2)} \tag{33}$$

The saddle points are then determined from

$$\frac{d}{d\alpha}[\xi\cos\alpha + \nu\alpha] = 0$$

which gives

$$\sin\alpha = \frac{\nu}{\xi}$$

There are two saddle points within the interval $-\pi/2$ and $3\pi/2$. For $\nu < \xi$, they are

$$\alpha = \sin^{-1}\frac{\nu}{\xi} \qquad \text{and} \qquad \alpha = \pi - \sin^{-1}\frac{\nu}{\xi} \tag{34}$$

with $0 \le \sin^{-1}(\nu/\xi) \le \pi/2$. For $\nu > \xi$, they are

$$\alpha = \frac{\pi}{2} + i\cosh^{-1}\frac{\nu}{\xi} \qquad \text{and} \qquad \alpha = \frac{\pi}{2} - i\cosh^{-1}\frac{\nu}{\xi} \tag{35}$$

In accordance with the original path defined for $H_\nu^{(1)}(\xi)$, the contributing saddle points are $\alpha = \sin^{-1}(\nu/\xi)$ for $\nu < \xi$ and $\alpha = \pi/2 - i\cosh^{-1}(\nu/\xi)$ for $\nu > \xi$.

Since there are no other singularities on the complex α plane, we can deform to the steepest descent path and carry out the saddle-point

method to find asymptotic series for $H_\nu^{(1)}(\xi)$. We can make use of (30) with the following identifications:

$$F(\alpha) = \frac{1}{\pi} e^{-i\nu\pi/2} \tag{36}$$

$$\xi f(\alpha) = i\xi \cos\alpha + i\nu\alpha \tag{37}$$

The leading terms in the asymptotic series are

$$H_\nu^{(1)}(\xi) \approx \sqrt{\frac{2}{\pi(\xi^2-\nu^2)^{1/2}}} \exp\left[i\left(\sqrt{\xi^2-\nu^2} - \nu\cos^{-1}\frac{\nu}{\xi} - \frac{\pi}{4}\right)\right] \tag{38}$$

for $\nu < \xi$ and

$$H_\nu^{(1)}(\xi) \approx -i\sqrt{\frac{2}{\pi(\nu^2-\xi^2)^{1/2}}} \exp\left[-\sqrt{\nu^2-\xi^2} + \nu\cosh^{-1}\frac{\nu}{\xi}\right] \tag{39}$$

for $\nu > \xi$. For the case $\nu \approx \xi$, the two contributing saddle points begin to merge into one and we need to evaluate the contribution due to both saddle points. It is readily seen that (38) reduces to (20) for $\nu \ll \xi$.

4.8 Integral Formulations for Dipoles in Layered Media

The geometrical configuration of the problem is shown in Figure 4.8.1. The origin of the coordinate system is placed in the location of the dipole which can be a vertical magnetic dipole (VMD), a vertical electric dipole (VED), a horizontal electric dipole (HED), or a horizontal magnetic dipole (HMD). There are M layers above the dipole at $z = d_1, d_2, \ldots d_M$ and N layers below it at $z = d_0, d_{-1}, \ldots, d_{-(N-1)}$. We shall first assume that all regions contain isotropic media. In region l, we denote the permittivity and permeability by ϵ_l and μ_l. Notice that in region $0, \epsilon_0$, and μ_0 are not necessarily equal to the free space permittivity and permeability which we denote by ϵ_o and μ_o.

The solutions to the wave equations can be written as superpositions of TE and TM wave components. Let A_l and B_l denote amplitudes for the TM waves and C_l and D_l denote amplitudes for the TE

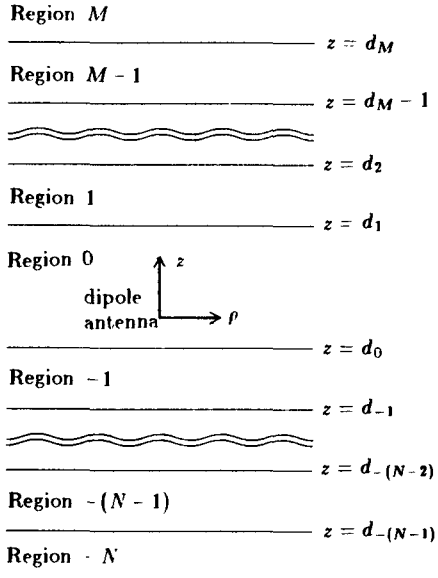


Figure 4.8.1 Dipole in layered medium.

waves. We find in region l the following solutions:

$$E_{lz} = \int_{-\infty}^{\infty} dk_\rho \left[A_l e^{ik_{lz}z} + B_l e^{-ik_{lz}z}\right] H_n^{(1)\prime}(k_\rho\rho) C_n(\phi) \tag{1}$$

$$E_{l\rho} = \int_{-\infty}^{\infty} dk_\rho \frac{ik_{lz}}{k_\rho} \left[A_l e^{ik_{lz}z} - B_l e^{-ik_{lz}z}\right] H_n^{(1)\prime}(k_\rho\rho) C_n(\phi)$$
$$+ \int_{-\infty}^{\infty} dk_\rho \frac{i\omega\mu_l}{k_\rho^2\rho} \left[C_l e^{ik_{lz}z} + D_l e^{-ik_{lz}z}\right] H_n^{(1)}(k_\rho\rho) S_n'(\phi) \tag{2}$$

$$E_{l\phi} = \int_{-\infty}^{\infty} dk_\rho \frac{ik_{lz}}{k_\rho^2\rho} \left[A_l e^{ik_{lz}z} - B_l e^{-ik_{lz}z}\right] H_n^{(1)}(k_\rho\rho) C_n'(\phi)$$
$$+ \int_{-\infty}^{\infty} dk_\rho \frac{-i\omega\mu_l}{k_\rho} \left[C_l e^{ik_{lz}z} + D_l e^{-ik_{lz}z}\right] H_n^{(1)\prime}(k_\rho\rho) S_n(\phi) \tag{3}$$

$$H_{lz} = \int_{-\infty}^{\infty} dk_\rho \left[C_l e^{ik_{lz}z} + D_l e^{-ik_{lz}z}\right] H_n^{(1)}(k_\rho\rho) S_n(\phi) \tag{4}$$

$$H_{l\rho} = \int_{-\infty}^{\infty} dk_\rho \frac{ik_{lz}}{k_\rho} \left[C_l e^{ik_{lz}z} - D_l e^{-ik_{lz}z}\right] H_n^{(1)\prime}(k_\rho\rho) S_n(\phi)$$

$$+\int_{-\infty}^{\infty} dk_\rho \frac{-i\omega\epsilon_l}{k_\rho^2\rho}\left[A_l e^{ik_{lz}z}+B_l e^{-ik_{lz}z}\right]H^{(1)}(k_\rho\rho)C_n'(\phi) \quad (5)$$

$$H_{l\phi}=\int_{-\infty}^{\infty} dk_\rho \frac{ik_{lz}}{k_\rho^2\rho}\left[C_l e^{ik_{lz}z}-D_l e^{-ik_{lz}z}\right]H_n^{(1)}(k_\rho\rho)S_n'(\phi)$$
$$+\int_{-\infty}^{\infty} dk_\rho \frac{i\omega\epsilon_l}{k_\rho}\left[A_l e^{ik_{lz}z}+B_l e^{-ik_{lz}z}\right]H^{(1)\prime}(k_\rho\rho)C_n(\phi) \quad (6)$$

In (1)–(6), $H_n^{(1)}(k_\rho\rho)$ is the nth order Hankel function of the first kind and $H_n^{(1)\prime}(k_\rho\rho)$ denotes the derivative of $H_n^{(1)}(\xi)$ with respect to its argument ξ. The ϕ-dependent functions $S_n(\phi)$ and $C_n(\phi)$ and the order of the Hankel functions n are all determined by the dipole configurations. The integrands of transverse field components $\overline{E}_s=\hat{\rho}E_\rho+\hat{\phi}E_\phi$ and $\overline{H}_s=\hat{\rho}H_\rho+\hat{\phi}H_\phi$ are derived from those of the longitudinal components E_z and H_z. Let

$$E_z=\int_{-\infty}^{\infty} dk_\rho\, E_z(k_\rho) \quad (7)$$

$$H_z=\int_{-\infty}^{\infty} dk_\rho\, H_z(k_\rho) \quad (8)$$

We have from Maxwell's equations in source-free regions

$$\overline{E}_s(k_\rho)=\frac{1}{k_\rho^2}\left[\nabla_s\frac{\partial}{\partial z}E_z(k_\rho)+i\omega\mu_l\nabla_s\times\overline{H}_z(k_\rho)\right] \quad (9)$$

$$\overline{H}_s(k_\rho)=\frac{1}{k_\rho^2}\left[\nabla_s\frac{\partial}{\partial z}H_z(k_\rho)-i\omega\epsilon_l\nabla_s\times\overline{E}_z(k_\rho)\right] \quad (10)$$

The boundary conditions of the interfaces require that tangential electric and magnetic field components be continuous for all ρ and ϕ. At $z=d_l$, we obtain

$$k_{lz}\left(A_l e^{ik_{lz}d_l}-B_l e^{-ik_{lz}d_l}\right)$$
$$=k_{(l-1)z}\left(A_{l-1}e^{ik_{(l-1)z}d_l}-B_{l-1}e^{-ik_{(l-1)z}d_l}\right) \quad (11)$$

$$\epsilon_l\left(A_l e^{ik_{lz}d_l}+B_l e^{-ik_{lz}d_l}\right)$$
$$=\epsilon_{(l-1)}\left(A_{l-1}e^{ik_{(l-1)z}d_l}+B_{l-1}e^{-ik_{(l-1)z}d_l}\right) \quad (12)$$

$$k_{lz}\left(C_l e^{ik_{lz}d_l} - D_l e^{-ik_{lz}d_l}\right) = k_{(l-1)z}\left(C_{l-1}e^{ik_{(l-1)z}d_l} - D_{l-1}e^{-ik_{(l-1)z}d_l}\right) \tag{13}$$

$$\mu_l\left(C_l e^{ik_{lz}d_l} + D_l e^{-ik_{lz}d_l}\right) = \mu_{(l-1)}\left(C_{l-1}e^{ik_{(l-1)z}d_l} + D_{l-1}e^{-ik_{(l-1)z}d_l}\right) \tag{14}$$

There are altogether $M+N$ boundaries which give rise to $4(M+N)$ equations as shown above. There are altogether $M+N+1$ regions. In regions M and $-N$ we have $A_M = C_M = 0$ and $B_{-N} = D_{-N} = 0$ because there are no waves originating from infinity. Thus we have a total of $4(M+N+1) - 4 = 4(M+N)$ unknowns to be solved from the $4(M+N)$ equations. The wave amplitudes are related to the configurations and the excitation amplitudes of the dipole antenna in region 0. The wave solutions in region 0 thus need special attention.

In the absence of the stratified medium, fields of a dipole radiating in unbounded space with permittivity ϵ_0 and μ_0 are well known. The solutions can be transformed from spherical coordinates to cylindrical coordinates by using the Sommerfeld identity

$$\frac{e^{ik_0 r}}{r} = \frac{i}{2}\int_{-\infty}^{\infty} dk_\rho \frac{k_\rho}{k_{0z}} H_0^{(1)}(k_\rho\rho)e^{ik_{0z}|z|} \tag{15}$$

In view of the four types of dipole configurations, we find that for

(1) Vertical electric dipole (VED)

$$E_z = \int_{-\infty}^{\infty} dk_\rho\, E_{ved}\begin{Bmatrix} e^{ik_{0z}z} \\ e^{-ik_{0z}z}\end{Bmatrix} H_0^{(1)}(k_\rho\rho) \qquad \begin{matrix} z \geq 0 \\ z \leq 0\end{matrix} \tag{16}$$

$$H_z = 0 \tag{17}$$

with

$$E_{ved} = -\frac{Ilk_\rho^3}{8\pi\omega\epsilon_0 k_{0z}} \tag{18}$$

where Il is the electric dipole moment.

(2) Horizontal electric dipole (HED)

$$E_z = \int_{-\infty}^{\infty} dk_\rho \begin{Bmatrix} E_{hed}\, e^{ik_{0z}z} \\ -E_{hed}\, e^{-ik_{0z}z}\end{Bmatrix} H_1^{(1)}(k_\rho\rho)\cos\phi \qquad \begin{matrix} z \geq 0 \\ z \leq 0\end{matrix} \tag{19}$$

$$H_z = \int_{-\infty}^{\infty} dk_\rho\, H_{hed}\begin{Bmatrix} e^{ik_{0z}z} \\ e^{-ik_{0z}z}\end{Bmatrix} H_1^{(1)}(k_\rho\rho)\sin\phi \qquad \begin{matrix} z \geq 0 \\ z \leq 0\end{matrix} \tag{20}$$

with

$$E_{hed} = i\frac{Ilk_\rho^2}{8\pi\omega\epsilon_0} \tag{21}$$

$$H_{hed} = i\frac{Ilk_\rho^2}{8\pi k_{0z}} \tag{22}$$

(3) Vertical magnetic dipole (VMD)

$$E_z = 0 \tag{23}$$

$$H_z = \int_{-\infty}^{\infty} dk_\rho\, H_{vmd} \left\{ \begin{matrix} e^{ik_{0z}z} \\ e^{-ik_{0z}z} \end{matrix} \right\} H_0^{(1)}(k_\rho\rho) \qquad \begin{matrix} z \geq 0 \\ z \leq 0 \end{matrix} \tag{24}$$

with

$$H_{vmd} = -i\frac{IA\,k_\rho^3}{8\pi k_{oz}} \tag{25}$$

where IA is the magnetic dipole moment.

(4) Horizontal magnetic dipole (HMD)

$$E_z = \int_{-\infty}^{\infty} dk_\rho\, E_{hmd} \left\{ \begin{matrix} e^{ik_{0z}z} \\ e^{-ik_{0z}z} \end{matrix} \right\} H_1^{(1)}(k_\rho\rho)\sin\phi \qquad \begin{matrix} z \geq 0 \\ z \leq 0 \end{matrix} \tag{26}$$

$$H_z = \int_{-\infty}^{\infty} dk_\rho \left\{ \begin{matrix} H_{hmd}\, e^{ik_{0z}z} \\ -H_{hmd} e^{-ik_{0z}z} \end{matrix} \right\} H_1^{(1)}(k_\rho\rho)\cos\phi \qquad \begin{matrix} z \geq 0 \\ z \leq 0 \end{matrix} \tag{27}$$

with

$$E_{hmd} = \frac{IA\omega\mu_0 k_\rho^2}{8\pi k_{0z}} \tag{28}$$

$$H_{hmd} = -\frac{IAk_\rho^2}{8\pi} \tag{29}$$

Notice that the magnetic dipoles produce fields which are duals of those produced by the corresponding electric dipoles. The results for the magnetic dipoles can be obtained by the replacement $\overline{E} \to \overline{H}$, $\overline{H} \to -\overline{E}$, $\mu_0 \rightleftarrows \epsilon_0$, and $Il \to i\omega\mu_0 IA$. We note in particular that at $z = 0$, the following field components vanish:

(1) VED

$$E_\rho = 0 \tag{30}$$

(2) HED

$$E_z = H_\rho = H_\phi = 0 \tag{31}$$

(3) VMD

$$H_\rho = 0 \tag{32}$$

(4) HMD

$$H_z = E_\rho = E_\phi = 0 \tag{33}$$

This is seen from (9)–(10) and by noting from (15) that

$$\frac{\partial}{\partial z}\frac{e^{ik_0 r}}{r} = 0 \qquad \text{at} \quad z = 0 \tag{34}$$

In the presence of the stratified medium, we can write the fields in region 0 by identifying A_0 and B_0 according to the four types of dipoles and whether we have $z > 0$ or $z < 0$. It becomes necessary that we distinguish the wave amplitudes in region 0 for $z \geq 0$ from those in region 0 for $z < 0$. For $z > 0$ we use A_{0+}, B_{0+}, C_{0+}, and D_{0+}; and for $z < 0$ we use A_{0-}, B_{0-}, C_{0-}, and D_{0-}. It is seen that for

(1) VED

$$\left.\begin{aligned} A_{0+} &= A_{ved} + E_{ved} & A_{0-} &= A_{ved} \\ B_{0+} &= B_{ved} & B_{0-} &= B_{ved} + E_{ved} \\ C_{0+} &= D_{0+} = C_{0-} = D_{0-} = 0 \end{aligned}\right\} \tag{35}$$

where A_{ved} and B_{ved} characterize contributions due to the stratified medium and are to be determined by the boundary conditions.

(2) HED

$$\left.\begin{aligned} A_{0+} &= A_{hed} + E_{hed} & A_{0-} &= A_{hed} \\ B_{0+} &= B_{hed} & B_{0-} &= B_{hed} - E_{hed} \\ C_{0+} &= C_{hed} + H_{hed} & C_{0-} &= C_{hed} \\ D_{0+} &= D_{hed} & D_{0-} &= D_{hed} + H_{hed} \end{aligned}\right\} \tag{36}$$

where A_{hed}, B_{hed}, C_{hed}, and D_{hed} characterize contributions due to the stratified medium and are to be determined by the boundary conditions.

(3) VMD

$$\left.\begin{aligned} &A_{0+} = B_{0+} = A_{0-} = B_{0-} = 0 \\ &C_{0+} = C_{vmd} + H_{vmd} \qquad C_{0-} = C_{vmd} \\ &D_{0+} = D_{vmd} \qquad D_{0-} = D_{vmd} + H_{vmd} \end{aligned}\right\} \tag{37}$$

where C_{vmd} and D_{vmd} characterize contributions due to the stratified medium and are to be characterized by the boundary conditions.

(4) HMD

$$\left.\begin{aligned} &A_{0+} = A_{hmd} + E_{hmd} \qquad A_{0-} = A_{hmd} \\ &B_{0+} = B_{hmd} \qquad B_{0-} = B_{hmd} + E_{hmd} \\ &C_{0+} = C_{hmd} + H_{hmd} \qquad C_{0-} = C_{hmd} \\ &D_{0+} = D_{hmd} \qquad D_{0-} = D_{hmd} - H_{hmd} \end{aligned}\right\} \tag{38}$$

where A_{hmd}, B_{hmd}, C_{hmd}, and D_{hmd} characterize contributions due to the stratified medium and are to be determined by the boundary conditions. We have expressed the solution in region 0 where the dipoles are located in terms of superpositions of the primary excitations in the absence of the stratified medium and the homogeneous solutions of the stratified medium in the absence of the source. The total wave amplitudes in region 0 are identified as A_{0+}, B_{0+}, C_{0+}, and D_{0+}, for $z \geq 0$ and A_{0-}, B_{0-}, C_{0-}, and D_{0-} for $z \leq 0$. It is easily shown that they satisfy the boundary conditions at $z = 0$ by remembering the vanishing field components as listed in (30)–(33) for the primary excitations.

We now determine the wave amplitudes in region 0. For TM waves, (11)–(12) can be solved to express A_l and B_l in terms of A_{l-1} and B_{l-1} or to express A_{l-1} and B_{l-1} in terms of A_l and B_l. We find

$$A_l e^{ik_{lz}d_l} = \frac{1}{2}\left[\frac{\epsilon_{l-1}}{\epsilon_l} + \frac{k_{(l-1)z}}{k_{lz}}\right] \cdot \left[A_{l-1}e^{ik_{(l-1)z}d_l} + R^{\mathrm{TM}}_{l(l-1)}B_{l-1}e^{-ik_{(l-1)z}d_l}\right] \tag{39a}$$

$$B_l e^{-ik_{lz}d_l} = \frac{1}{2}\left[\frac{\epsilon_{l-1}}{\epsilon_l} + \frac{k_{(l-1)z}}{k_{lz}}\right] \cdot \left[R^{\mathrm{TM}}_{l(l-1)}A_{l-1}e^{ik_{(l-1)z}d_l} + B_{l-1}e^{-ik_{(l-1)z}d_l}\right] \tag{39b}$$

for A_l and B_l in terms of A_{l-1} and B_{l-1}, and

$$A_{l-1}e^{ik_{(l-1)z}d_l} = \frac{1}{2}\left(\frac{\epsilon_l}{\epsilon_{l-1}} + \frac{k_{lz}}{k_{(l-1)z}}\right) \cdot \left[A_l e^{ik_{lz}d_l} + R^{\mathrm{TM}}_{(l-1)l} B_l e^{-ik_{lz}d_l}\right] \tag{40a}$$

$$B_{l-1}e^{-ik_{(l-1)z}d_l} = \frac{1}{2}\left(\frac{\epsilon_l}{\epsilon_{l-1}} + \frac{k_{lz}}{k_{(l-1)z}}\right) \cdot \left[R^{\mathrm{TM}}_{(l-1)l} A_l e^{ik_{lz}d_l} + B_l e^{-ik_{lz}d_l}\right] \tag{40b}$$

for A_{l-1} and B_{l-1} in terms of A_l and B_l. In (39)–(40), the Fresnel reflection coefficient

$$R^{\mathrm{TM}}_{(l-1)l} = \frac{1 - \epsilon_{l-1}k_{lz}/\epsilon_l k_{(l-1)z}}{1 + \epsilon_{l-1}k_{lz}/\epsilon_l k_{(l-1)z}} = -R^{\mathrm{TM}}_{l(l-1)} \tag{41}$$

A similar procedure applies to the case of TE waves. The results are duals of those of (39)–(41) with the replacements of A by C, B by D, and ϵ by μ.

For $z \geq 0$, we notice that $B_M = D_M = 0$. Letting $l = 0$, we obtain the reflection coefficients $R^{\mathrm{TM}}_{0+} = B_{0+}/A_{0+}$ and $R^{\mathrm{TE}}_{0+} = D_{0+}/C_{0+}$ in the form of continued fractions. We find

$$R^{\mathrm{TM}}_{0+} = \frac{B_{0+}}{A_{0+}} = \frac{e^{i2k_{0z}d_1}}{R^{\mathrm{TM}}_{01}} + \frac{\left[1 - \left(1/R^{\mathrm{TM}}_{01}\right)^2\right] e^{i2(k_{0z}+k_{1z})d_1}}{(1/R^{\mathrm{TM}}_{01})e^{i2k_{1z}d_1} + (B_1/A_1)} \tag{42}$$

$$R^{\mathrm{TE}}_{0+} = \frac{D_{0+}}{C_{0+}} = \frac{e^{i2k_{0z}d_1}}{R^{\mathrm{TE}}_{01}} + \frac{\left[1 - \left(1/R^{\mathrm{TE}}_{01}\right)^2\right] e^{i2(k_{0z}+k_{1z})d_1}}{(1/R^{\mathrm{TE}}_{01})e^{i2k_{1z}d_1} + (D_1/C_1)} \tag{43}$$

where B_1/A_1 and D_1/C_1 can be expressed in terms of B_2/A_2 and D_2/C_2 and so on until region M where $B_M/A_M = 0 = D_M/C_M$.

For $z \leq 0$, we notice that $A_{-N} = C_{-N} = 0$. Letting $l = 0$, we obtain the reflection coefficients $R^{\mathrm{TM}}_{0-} = A_{0-}/B_{0-}$ and $R^{\mathrm{TE}}_{0-} = C_{0-}/D_{0-}$ in the form of continued fractions. We find

$$R^{\mathrm{TM}}_{0-} = \frac{A_{0-}}{B_{0-}} = \frac{e^{-i2k_{0z}d_0}}{R^{\mathrm{TM}}_{0(-1)}} + \frac{\left[1 - \left(1/R^{\mathrm{TM}}_{0(-1)}\right)^2\right] e^{-i2(k_{0z}+k_{-1z})d_0}}{(1/R^{\mathrm{TM}}_{0(-1)})e^{-i2k_{-1z}d_0} + (A_{-1}/B_{-1})} \tag{44}$$

$$R_{0-}^{\mathrm{TE}} = \frac{C_{0-}}{D_{0-}}$$

$$= \frac{e^{-i2k_{0z}d_0}}{R_{0(-1)}^{\mathrm{TE}}} + \frac{\left[1 - \left(1/R_{0(-1)}^{\mathrm{TE}}\right)^2\right] e^{-i2(k_{0z}+k_{-1z})d_0}}{\left(1/R_{0(-1)}^{\mathrm{TE}}\right) e^{-i2k_{-1z}d_0} + (C_{-1}/D_{-1})} \tag{45}$$

where A_{-1}/B_{-1} and C_{-1}/D_{-1} are expressible in terms of A_{-2}/B_{-2} and C_{-2}/D_{-2} and so on until region $-N$, where $A_{-N}/B_{-N} = 0$ and $C_{-N}/D_{-N} = 0$.

Once the wave amplitudes in region 0 are found, wave amplitudes in other regions can be determined by the use of propagation matrices, which are readily determined from (39) and (40) and from a set of dual equations for TE waves. We now determine the wave amplitudes in region 0.

(1) VED: From (35), we find

$$R_{0+}^{\mathrm{TM}} = \frac{B_{0+}}{A_{0+}} = \frac{B_{ved}}{A_{ved} + E_{ved}} \tag{46a}$$

$$R_{0-}^{\mathrm{TM}} = \frac{A_{0-}}{B_{0-}} = \frac{A_{ved}}{B_{ved} + E_{ved}} \tag{46b}$$

Solving from A_{ved} and B_{ved} and substituting in (35), we obtain

$$A_{0+} = \frac{1 + R_{0-}^{\mathrm{TM}}}{1 - R_{0+}^{\mathrm{TM}} R_{0-}^{\mathrm{TM}}} E_{ved} \tag{47a}$$

$$B_{0+} = \frac{R_{0+}^{\mathrm{TM}}(1 + R_{0-}^{\mathrm{TM}})}{1 - R_{0+}^{\mathrm{TM}} R_{0-}^{\mathrm{TM}}} E_{ved} \tag{47b}$$

$$A_{0-} = \frac{R_{0-}^{\mathrm{TM}}(1 + R_{0+}^{\mathrm{TM}})}{1 - R_{0+}^{\mathrm{TM}} R_{0-}^{\mathrm{TM}}} E_{ved} \tag{47c}$$

$$B_{0-} = \frac{1 + R_{0+}^{\mathrm{TM}}}{1 - R_{0+}^{\mathrm{TM}} R_{0-}^{\mathrm{TM}}} E_{ved} \tag{47d}$$

(2) HED: By the same token, we find from (36)

$$A_{0+} = \frac{1 - R_{0-}^{\mathrm{TM}}}{1 - R_{0+}^{\mathrm{TM}} R_{0-}^{\mathrm{TM}}} E_{hed} \tag{48a}$$

$$B_{0+} = \frac{R_{0+}^{\mathrm{TM}}(1 - R_{0-}^{\mathrm{TM}})}{1 - R_{0+}^{\mathrm{TM}} R_{0-}^{\mathrm{TM}}} E_{hed} \tag{48b}$$

$$C_{0+} = \frac{1 + R_{0-}^{\mathrm{TE}}}{1 - R_{0+}^{\mathrm{TE}} R_{0-}^{\mathrm{TE}}} H_{hed} \tag{48c}$$

$$D_{0+} = \frac{R_{0+}^{\mathrm{TE}}(1 + R_{0-}^{\mathrm{TE}})}{1 - R_{0+}^{\mathrm{TE}} R_{0-}^{\mathrm{TE}}} H_{hed} \tag{48d}$$

$$A_{0-} = -\frac{R_{0-}^{\mathrm{TM}}(1 - R_{0+}^{\mathrm{TM}})}{1 - R_{0+}^{\mathrm{TM}} R_{0-}^{\mathrm{TM}}} E_{hed} \tag{49a}$$

$$B_{0-} = -\frac{1 - R_{0+}^{\mathrm{TM}}}{1 - R_{0+}^{\mathrm{TM}} R_{0-}^{\mathrm{TM}}} E_{hed} \tag{49b}$$

$$C_{0-} = \frac{R_{0-}^{\mathrm{TE}}(1 + R_{0+}^{\mathrm{TE}})}{1 - R_{0+}^{\mathrm{TE}} R_{0-}^{\mathrm{TE}}} H_{hed} \tag{49c}$$

$$D_{0-} = \frac{1 + R_{0+}^{\mathrm{TE}}}{1 - R_{0+}^{\mathrm{TE}} R_{0-}^{\mathrm{TE}}} H_{hed} \tag{49d}$$

(3) VMD: From (37), we find

$$C_{0+} = \frac{1 + R_{0-}^{\mathrm{TE}}}{1 - R_{0+}^{\mathrm{TE}} R_{0-}^{\mathrm{TE}}} H_{vmd} \tag{50a}$$

$$D_{0+} = \frac{R_{0+}^{\mathrm{TE}}(1 + R_{0-}^{\mathrm{TE}})}{1 - R_{0+}^{\mathrm{TE}} R_{0-}^{\mathrm{TE}}} H_{vmd} \tag{50b}$$

$$C_{0-} = \frac{R_{0-}^{\mathrm{TE}}(1 + R_{0+}^{\mathrm{TE}})}{1 - R_{0+}^{\mathrm{TE}} R_{0-}^{\mathrm{TE}}} H_{vmd} \tag{50c}$$

$$D_{0-} = \frac{1 + R_{0+}^{\mathrm{TE}}}{1 - R_{0+}^{\mathrm{TE}} R_{0-}^{\mathrm{TE}}} H_{vmd} \tag{50d}$$

(4) HMD: From (38), we find

$$A_{0+} = \frac{1 + R_{0-}^{\mathrm{TM}}}{1 - R_{0+}^{\mathrm{TM}} R_{0-}^{\mathrm{TM}}} E_{hmd} \tag{51a}$$

$$B_{0+} = \frac{R_{0+}^{\mathrm{TM}}(1 + R_{0-}^{\mathrm{TM}})}{1 - R_{0+}^{\mathrm{TM}} R_{0-}^{\mathrm{TM}}} E_{hmd} \tag{51b}$$

$$C_{0+} = \frac{1 - R_{0-}^{\mathrm{TE}}}{1 - R_{0+}^{\mathrm{TE}} R_{0-}^{\mathrm{TE}}} H_{hmd} \tag{51c}$$

$$D_{0+} = \frac{R_{0+}^{\mathrm{TE}}(1 - R_{0-}^{\mathrm{TE}})}{1 - R_{0+}^{\mathrm{TE}} R_{0-}^{\mathrm{TE}}} H_{hmd} \tag{51d}$$

$$A_{0-} = \frac{R_{0-}^{\mathrm{TM}}(1 + R_{0+}^{\mathrm{TM}})}{1 - R_{0+}^{\mathrm{TM}} R_{0-}^{\mathrm{TM}}} E_{hmd} \tag{52a}$$

$$B_{0-} = \frac{1 + R_{0+}^{\mathrm{TM}}}{1 - R_{0+}^{\mathrm{TM}} R_{0-}^{\mathrm{TM}}} E_{hmd} \tag{52b}$$

$$C_{0-} = -\frac{R_{0-}^{\mathrm{TE}}(1 - R_{0+}^{\mathrm{TE}})}{1 - R_{0+}^{\mathrm{TE}} R_{0-}^{\mathrm{TE}}} H_{hmd} \tag{52c}$$

$$D_{0-} = -\frac{1 - R_{0+}^{\mathrm{TE}}}{1 - R_{0+}^{\mathrm{TE}} R_{0-}^{\mathrm{TE}}} H_{hmd} \tag{52d}$$

The solutions for the electromagnetic field components are obtained by inserting the values for the wave amplitudes into (1)–(6).

4.9 Dipole on One-Layer Medium

Consider a vertical magnetic dipole (VMD) placed on the surface of a one-layer (or half-space) medium with permeability μ and permittivity ϵ_1 [Fig. 4.9.1]. The vertical magnetic field component H_z as measured by an observer in region zero where $\mu_0 = \mu$ and $\epsilon_0 = \epsilon$, takes the form

$$H_z = -i\frac{IA}{8\pi} \int_{SIP} dk_\rho \frac{k_\rho^3}{k_z}(1 + R_{01}^{\mathrm{TE}}) H_0^{(1)}(k_\rho \rho) e^{ik_z z} \tag{1}$$

where

$$1 + R_{01}^{\mathrm{TE}} = \frac{2k_z}{k_z + k_{1z}} \tag{2}$$

$$k_z = \sqrt{k^2 - k_\rho^2}$$

$$k_{1z} = \sqrt{k_1^2 - k_\rho^2}$$

$$k^2 = \omega^2 \mu \epsilon$$

$$k_1^2 = \omega^2 \mu \epsilon_1$$

The Sommerfeld integration path (SIP) runs from $-\infty$ to ∞ in such a manner that it is slightly above the negative real $k_{\rho R}$ axis for $-\infty <$

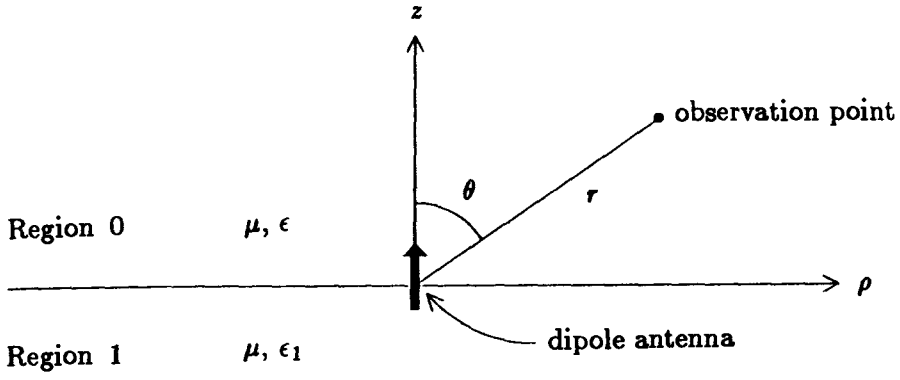


Figure 4.9.1 Dipole on one-layer medium.

$k_{\rho R} \leq 0$ and slightly below the positive real $k_{\rho R}$ axis for $0 < k_{\rho R} < \infty$ [Fig. 4.9.2].

We apply the saddle-point method to evaluate the integral (1) for very large observation distance ρ and z with $\rho = r \sin\theta$ and $z = r \cos\theta$. Substituting (2) in (1) and making use of the asymptotic form for the Hankel function

$$H_0^{(1)}(k_\rho \rho) \approx \sqrt{2/i\pi k_\rho \rho}\, e^{ik_\rho \rho} \tag{3}$$

we obtain

$$H_z = -i\frac{IA}{4\pi}\sqrt{\frac{2}{i\pi\rho}} \int_{\text{SIP}} dk_\rho\, F(k_\rho) e^{i\rho f(k_\rho)} \tag{4}$$

where

$$F(k_\rho) = \frac{k_\rho^{5/2}}{k_z + k_{1z}} \tag{5}$$

and

$$f(k_\rho) = k_\rho + k_z \cot\theta \tag{6}$$

First, we find the saddle point by setting $f'(k_\rho) = 0$ which yields

$$k_\rho = k \sin\theta \tag{7}$$

We do not concern ourselves with the other saddle point $k_\rho = -k \sin\theta$ as we shall deform the SIP to the steepest descent path (SDP) on the upper k_ρ plane where $e^{ik_\rho \rho}$ converges.

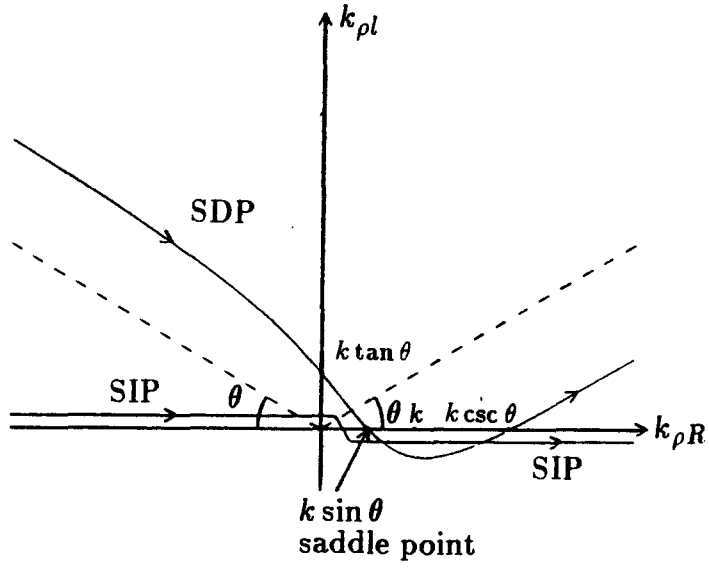


Figure 4.9.2 SIP and SDP on complex k plane.

We then make use of the result from the saddle-point method, which gives

$$H_z = -\frac{IA}{2\pi\rho}\sqrt{-1/f''}\,F(k_\rho = k\sin\theta)e^{if(k_\rho = k\sin\theta)} \left\{1 + \frac{1}{i2\rho f''}\left[\frac{f'''}{f''}\frac{F'}{F} + \frac{1}{4}\frac{f^{iv}}{f''} - \frac{5}{12}\left(\frac{f'''}{f''}\right)^2 - \frac{F''}{F}\right] + \cdots\right\} \tag{8}$$

The first term yields results to the order of $1/r$. From (6), we find

$$f''(k_\rho) = -k^2\cot\theta/k_z^3 \tag{9}$$

At the saddle point

$$f(k_\rho = k\sin\theta) = k/\sin\theta \tag{10a}$$

$$f''(k_\rho = k\sin\theta) = -1/k\sin\theta\cos^2\theta \tag{10b}$$

$$F(k_\rho = k\sin\theta) = \frac{k^{5/2}\sin^{5/2}\theta}{k\cos\theta + \sqrt{k_1^2 - k^2\sin^2\theta}} \tag{10c}$$

Thus we obtain from the first term of (8)

$$H_z = -\frac{IAe^{ikr}}{2\pi r}\frac{k^2\sin^2\theta\cos\theta}{\cos\theta + \sqrt{k_1^2/k^2 - \sin^2\theta}} \tag{11}$$

We notice that $H_z = 0$ at both $\theta = 0$ and $\theta = \pi/2$. At $\theta = 0$, the integrand in (4) is zero even before using the saddle-point method, which is expected physically because a vertical magnetic dipole should not give rise to a nonvanishing H_z along the axis of the VMD.

When the observation point is along the surface where $\theta = \pi/2$ and $z = 0$, we do not expect H_z to be identically equal to zero. Thus we should evaluate the saddle point contribution to second order of $(1/r)^2$ by making full use of (8). We obtain from (9) and (5),

$$f'''(k_\rho) = -3k^2 k_\rho \cot\theta / k_z^5 \tag{12a}$$

$$f^{iv}(k_\rho) = -3k^2 \cot\theta [k^2 + 4k_\rho^2]/k_z^7 \tag{12b}$$

$$F'(k_\rho) = \frac{5k_\rho^{3/2} k_z k_{1z} + 2k_\rho^{7/2}}{2k_z k_{1z}(k_z + k_{1z})} \tag{13a}$$

$$F''(k_\rho) = \Big\{ k_\rho^{5/2} \left[5k_z k_{1z} + 2k_\rho^2\right] \left[k_z^2 + k_z k_{1z} + k_{1z}^2\right] + k_\rho^{1/2} k_z k_{1z} \left[\frac{15}{2} k_z^2 k_{1z}^2 - 5k_\rho^2 (k_z^2 + k_{1z}^2) + 7k_\rho^2 k_z k_{1z}\right] \Big\} \Big/ 2k_z^3 k_{1z}^3 (k_z + k_{1z}) \tag{13b}$$

In order to evaluate (8) for $\theta = \pi/2$, which gives $k_z = 0$ and $k_\rho = k$, we notice that at the saddle point $k_\rho = k \sin\theta$, f'' and F have been given in (10*b*) and (10*c*). We shall keep the largest terms in (12) and (13) as θ approaches $\pi/2$. It follows that

$$f''' = -3/k^2 \cos^4\theta \tag{14a}$$

$$f^{iv} = -15/k^3 \cos^6\theta \tag{14b}$$

$$F' = k^{5/2}/k_{1z}^2 \cos\theta \tag{14c}$$

$$F'' = k^{3/2}/k_{1z}^2 \cos^3\theta \tag{14d}$$

Introducing (10) and (14) into (8) and setting $\theta = \pi/2$, we find from the saddle point contribution

$$H_z = i \frac{k^3 I A e^{ik\rho}}{2\pi(k_1^2 - k^2)\rho^2} \tag{15}$$

We see that on the surface of the half-space medium, the wave is propagating with the wavenumber k and the wave amplitude decays as

$1/\rho^2$ away from the VMD. Since the VMD is placed at the interface, by symmetry we expect that there should be a wave propagating with the wavenumber k_1. To look for this solution, we must fully understand all the singularities on the complex k_ρ plane and carefully evaluate their contributions when the SIP is detoured to the steepest descent path (SDP) by following the saddle-point method.

Careful examination of (1) shows that there are branch points at $k_\rho = \pm k$ due to k_z, at $k_\rho = \pm k_1$ due to k_{1z}, and at $k_\rho = 0$ due to $H_0^{(1)}(k_\rho \rho)$. Before choosing the branch cuts, we first locate the SDP by demanding that the real part of $f(k_\rho)$ in (6) be equal to that at the saddle point so that the exponential phase term in (4) does not vary along the SDP,

$$Re\{k_\rho + k_z \cot\theta\} = k/\sin\theta \tag{16a}$$

Furthermore, the convergence requirement demands that

$$Im\{k_\rho + k_z \cot\theta\} \geq 0 \tag{16b}$$

The asymptotics as k_ρ approaches infinity are determined from (16) by letting $k_{\rho R} \to \infty$ and $k_{\rho I} \to \infty$ which gives

$$k_{\rho R} \mp k_{\rho I} \cot\theta = k/\sin\theta \tag{17a}$$

$$k_{\rho I} \pm k_{\rho R} \cot\theta \geq 0 \tag{17b}$$

From (17b) we see that for $k_{\rho R} \to -\infty$ we choose the bottom sign and for $k_{\rho R} \to \infty$ we choose the upper sign. Equation (17a) then yields

$$\frac{k_{\rho I}}{k_{\rho R}} = \pm\tan\theta \qquad \text{as} \qquad k_{\rho R} \to \pm\infty \tag{18a}$$

The asymptotics are plotted in Figure 4.9.2.

We now determine the intersections with the real $k_{\rho R}$ axis and the imaginary $k_{\rho I}$ axis for the SDP. Setting $k_{\rho R} = 0$ in (16a) we find

$$\cot\theta\, Re\left\{\sqrt{k^2 - (ik_{\rho I})^2}\right\} = k/\sin\theta$$

which gives

$$k_{\rho I} = k\tan\theta \tag{18b}$$

Setting $k_{\rho I} = 0$ in (16a) we find

$$k_{\rho R} + \cot\theta\, Re\left\{\sqrt{k^2 - k_{\rho R}^2}\right\} = k/\sin\theta$$

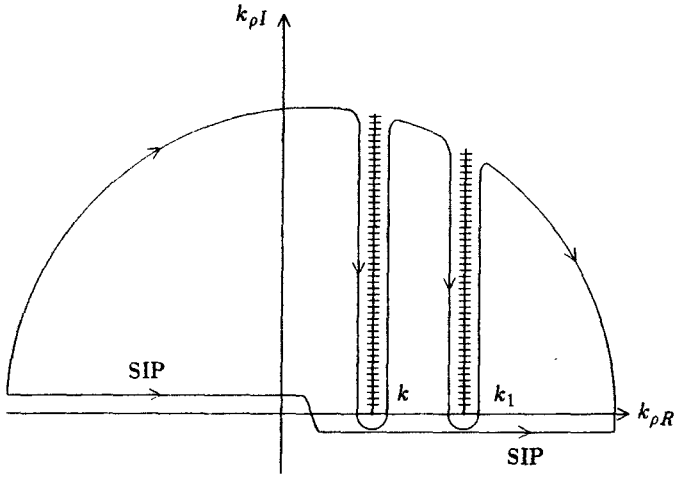


Figure 4.9.3 Vertical branch cuts.

which gives, for $k_{\rho R} > k$

$$k_{\rho R} = k \csc\theta \tag{18c}$$

and for $k_{\rho R} < k$

$$k_{\rho R} = k \sin\theta \tag{18d}$$

Clearly (18*d*) indicates that the SDP passes through the saddle point. The SDP is shown in Figure 4.9.2.

As $\theta = \pi/2$, we see from Figure 4.9.2 that the SDP becomes a vertical path going downward on the left-hand side of $k_\rho = k$ and upward on the right-hand side of $k_\rho = k$, where $k_\rho = k$ is now the saddle point. We choose both the branch cuts originating from $k_\rho = k$ and $k_\rho = k_1$ to be vertical. When the original integration path SIP is deformed to the SDP passing around $k_\rho = k$, it also passes around $k_\rho = k_1$ in the same manner [Fig. 4.9.3]. In fact we see from (1) and (2) that as $z = 0$ (namely $\theta = \pi/2$), the integral is symmetric with respect to k_z and k_{1z}. Thus by parallel analysis, the branch cut contribution from k_1 gives results similar to those obtained from the saddle-point method applied to the fields in Region 1. At the interface, we conclude that

$$H_z = i\frac{IA}{2\pi(k_1^2 - k^2)\rho^2}\left\{k^3 e^{ik\rho} - k_1^3 e^{ik_1\rho}\right\} \tag{19}$$

where the second term is obtained by interchanging k_1 and k in the first term. Therefore, at the interface there are two waves propagating with wavenumbers k and k_1 and decaying as $1/\rho^2$ with distance ρ. As a function of distance along the surface, these two waves interfere and give rise to an interference pattern. The periodicity of the pattern is proportional to $2\pi/(k_1 - k)$.

4.10 Dipole Above Two-Layer Medium

Consider a dipole antenna located on the surface of a two-layer medium [Fig. 4.10.1]. The field H_z arising from a vertical magnetic dipole is, on the surface of the medium,

$$H_z = -i\frac{IA}{8\pi}\int_{-\infty}^{\infty} dk_\rho \frac{k_\rho^3}{k_z}[1 + R^{\mathrm{TE}}]H_0^{(1)}(k_\rho \rho)e^{ik_z z} \tag{1}$$

where

$$R^{\mathrm{TE}} = \frac{R_{01} + R_{12}e^{i2k_{1z}d}}{1 + R_{01}R_{12}e^{i2k_{1z}d}} \tag{2}$$

We evaluate H_z for the field on the surface and very far away from the source.

Since the magnitudes of R_{01} and R_{12} are less than unity, we make use of the fact that $R_{01} = -R_{10}$ and expand the denominator of the integrand:

$$\begin{aligned} 1 + R^{\mathrm{TE}} &= 1 + (-R_{10} + R_{12}e^{i2k_{1z}d})\sum_{m=0}^{\infty}(R_{10}R_{12}e^{i2k_{1z}d})^m \\ &= 1 - \sum_{m=0}^{\infty} R_{10}^{m+1}(R_{12}e^{i2k_{1z}d})^m + \sum_{m=1}^{\infty} R_{10}^{m-1}(R_{12}e^{i2k_{1z}d})^m \\ &= (1 - R_{10}) + \sum_{m=1}^{\infty}(1 - R_{10}^2)(R_{10}^{m-1}R_{12}^m e^{i2mk_{1z}d}) \end{aligned} \tag{3}$$

Substituting (3) in (1) we obtain

$$\begin{aligned} H_z = &-i\frac{IA}{8\pi}\int_{-\infty}^{\infty} dk_\rho \frac{k_\rho^3}{k_z}(1 + R_{01})H_0^{(1)}(k_\rho \rho) \\ &-i\frac{IA}{8\pi}\sum_{m=1}^{\infty}\int_{-\infty}^{\infty} dk_\rho \frac{k_\rho^3}{k_z}(1 - R_{10}^2)R_{10}^{m-1}R_{12}^m e^{ik_{1z}(2md)}H_0^{(1)}(k_\rho \rho) \end{aligned} \tag{4}$$

The first term has been evaluated and yields the half-space solution. The summation term accounts for contributions from the subsurface.

Each term in the summation term can be evaluated with the saddle-point method. Notice that the Hankel function has the asymptotic value, as $k_\rho \rho \to \infty$,

$$H_0^{(1)}(k_\rho \rho) \sim \sqrt{\frac{2}{i\pi k_\rho \rho}}\, e^{ik_\rho \rho} \tag{5}$$

First we make the transformation to the complex α plane

$$k_\rho = k_1 \sin\alpha \qquad k_{1z} = k_1 \cos\alpha \tag{6}$$

$$\rho = R_m \sin\alpha_m \qquad 2md = R_m \cos\alpha_m \tag{7}$$

where

$$R_m = \sqrt{\rho^2 + (2md)^2} \tag{8}$$

The integral for the m th term then becomes

$$H_{zm} = \int d\alpha\, G_m(\alpha) e^{ik_1 R_m \cos(\alpha - \alpha_m)} \tag{9}$$

where

$$G_m(\alpha) = -i\frac{IA}{8\pi}\frac{k_1 \cos\alpha k_1^3 \sin^3\alpha}{\sqrt{k^2 - k_1^2 \sin^2\alpha}} \cdot (1 - R_{10}^2) R_{10}^{m-1} R_{12}^m \sqrt{\frac{2}{i\pi k R_m \sin\alpha \sin\alpha_m}} \tag{10}$$

The integrand in (9) is dominated by the exponential term as $k_1 R_m \approx k_1 \rho \gg 1$ under the assumption of $k_1 z \ll 1$. A saddle-point occurs at $\alpha = \alpha_m$. The saddle point contribution is calculated to be, to first order in $1/R_m$,

$$H_{zm} = \sqrt{2\pi/ik_1 R_m}\, G_m(\alpha = \alpha_m) e^{ik_1 R_m} \sim \frac{1}{R_m} e^{ik_1 R_m} \tag{11}$$

Notice that $1/R_m$ results because $G_m(\alpha)$ is proportional to $R_m^{-1/2}$.

Thus at a point ρ on the surface the field is a superposition of the half-space solution H_{zh} plus a series of direct waves radiated by

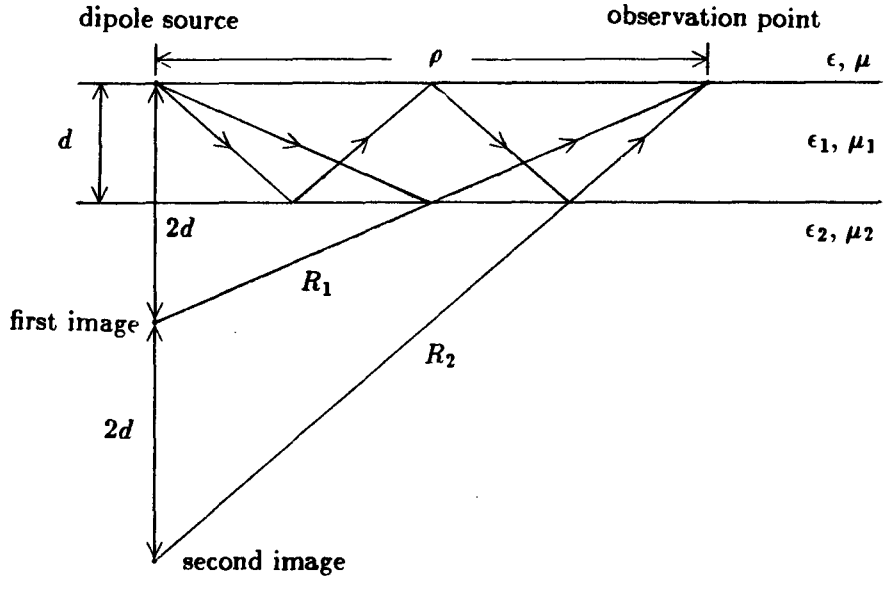


Figure 4.10.1 Dipole source and observation point on two layer medium.

sources located at $R_m = [\rho^2 + (2md)^2]^{1/2}$. These series of sources can be viewed as images of the original dipole located at distances $2md$ below the surface [Fig. 4.10.1]. It may also be viewed as being due to waves generated by the original dipole and reflected m times at the second boundary surface before reaching the observation point.

One application of this problem is in the area of geophysical subsurface probing. For a vertical magnetic dipole placed on a half-space medium, the direction of maximum power coupling into medium 1 is along the critical angle $\alpha_c = \sin^{-1}(k/k_1)$. On the surface of the medium, the field solution is proportional to $1/\rho^2$. For the vertical magnetic field component, we have

$$H_z = -\frac{iIA}{2\pi(k_1^2 - k^2)\rho^2}[k_1^3 e^{ik_1\rho} - k^3 e^{ik\rho}]$$

Since only field solutions proportional to $1/r$ give rise to radiated power, the radiation pattern has a null on the surface. The field solution on the surface is a superposition of two waves propagating at different phase velocities and giving rise to an interference pattern. When there is a subsurface, the interference pattern is modified due to the series of

image sources. Since the field is maximum in the direction of the critical angle, we expect the field to be reflected by the surface and contribute to a higher magnitude for H_z in the interference plot. The positions of such high magnitudes occur at $\rho_m = 2md\tan\alpha_c = 2mkd/(k_1^2-k^2)^{1/2}$.

We note that in this approach, which is known as the geometrical optics approximation, the summation series converges faster when the layer is thicker, corresponding to a larger d. The same is true when medium 1 is more lossy, that is, when k_{1z} has a larger imaginary part. As the layer thickness decreases, more terms in the geometrical optics approximation are needed in order to ensure convergence.

The normal mode approach becomes more convenient when the layer thickness is small. Instead of expanding (2), we find all poles from $1+R^{\mathrm{TE}}$ and express the integral (1) in terms of a residue series. Poles are determined by setting the denominator in (2) equal to zero. We write

$$R_{10}R_{12}e^{i2k_{1z}d} = e^{i2l\pi} \tag{12}$$

where l is any integer. Equation (12) is similar to that used in the calculation of guidance conditions for a slab waveguide. Each pole corresponds to a wave mode. For the poles that lie between the original path of integration SDP and the steepest descent path, the corresponding wave modes are excited. The saddle-point contribution corresponds to a direct wave from the transmitter to the receiver. For the poles that lie close to the saddle-point, their effect on the saddle-point contribution must be taken into account by using the modified saddle-point method [Problem P4.36].

PROBLEMS

Problem P4.1

An alternate approach to the solution of radiation problems is by means of vector potentials. This approach is especially useful for isotropic media. Difficulties will arise when radiation occurs in non-isotropic media. The vector potential $\overline{A}$ is defined as

$$\mu\overline{H} = \nabla\times\overline{A}$$

(a) Show that this definition does not uniquely determine $\overline{A}$, since a gauge transformation with $\overline{A}' = \overline{A}+\nabla\psi$, where ψ is any scalar

function, gives rise to the same $\overline{H}$ field. The additional condition, which specifies the divergence of $\overline{A}$, is known as the gauge condition. The Lorentz gauge condition is

$$\nabla \cdot \overline{A} - i\omega\mu\epsilon\phi = 0$$

where ϕ is a scalar potential in terms of which

$$\overline{E} = i\omega\overline{A} - \nabla\phi$$

(b) Show that with the Lorentz gauge, the equation for $\overline{A}$ becomes

$$(\nabla^2 + \omega^2\mu\epsilon)\overline{A} = -\mu\overline{J}$$

Using the scalar Green's function, determine $\overline{A}$ in terms of $\overline{J}$. With this approach, find the electric field vector $\overline{E}$. Show that the electric field vector $\overline{E}$, when expressed in terms of $\overline{A}$, yields results identical to (4.3.16).

Problem P4.2

For source distributions of a finite extent radiating in unbounded space, boundary conditions must be imposed at infinity to obtain unique solutions to the radiation problem. Such boundary conditions are called radiation conditions and require that solutions attenuate no slower than the inverse distance far away from the source and that the wave must propagate outward to infinity. In mathematical terms, the radiation conditions for $\overline{E}$ and $\overline{H}$ take the form

$$\lim_{r\to\infty} r[\overline{H} - \hat{r} \times \overline{E}/\eta] = 0$$

$$\lim_{r\to\infty} r[\overline{E} + \hat{r} \times \eta\overline{H}] = 0$$

(a) Show that these conditions are satisfied with the radiation fields

$$\overline{E} = i\omega\mu\frac{e^{ikr}}{4\pi r}(\hat{\theta}f_\theta + \hat{\phi}f_\phi)$$

$$\overline{H} = i\omega\mu\frac{e^{ikr}}{4\pi r}\eta(\hat{\phi}f_\theta - \hat{\theta}f_\phi)$$

where f_θ and f_ϕ are the $\hat{\theta}$ and $\hat{\phi}$ components of the vector current moment.

(b) Applying the Maxwell equations for $\overline{E}$ and $\overline{H}$, show that

$$\lim_{r\to\infty} r[\nabla \times \overline{E} - ik\hat{r} \times \overline{E}] = 0$$
$$\lim_{r\to\infty} r[\nabla \times \overline{H} - ik\hat{r} \times \overline{H}] = 0$$

and show that the radiation condition for the dyadic Green's function is

$$\lim_{r\to\infty} r[\nabla \times \overline{\overline{G}}(\overline{r},\overline{r}') - ik\hat{r} \times \overline{\overline{G}}(\overline{r},\overline{r}')] = 0$$

(c) Impose the radiation condition for scalar waves

$$\lim_{r\to\infty} r[\frac{\partial U}{\partial r} - ikU] = 0$$

and investigate the uniqueness theorem. The solutions are assumed to be composed of all spherical harmonic solutions of the scalar wave equation

$$(\nabla^2 + k^2)U = 0$$

Is the radiation condition at infinity and the boundary condition at the source sufficient in determining the amplitudes of all the spherical harmonics? Expressing all higher-order harmonics f_n with $n > 0$ in terms of the fundamental f_0 with recurrence relations and show that the radiation condition uniquely determines f_0. Since f_0 is an outgoing wave so are all the higher-order harmonics.

Problem P4.3

(a) Find the electric and magnetic fields due to a line current source with $I(z) = I_0 e^{ik_z z}$, placed along the z axis in free space.
(b) Evaluate the real part of the complex Poynting's vector in the far field. What happens if $k_z > k$?
(c) Sketch the shape of the equi-phase surfaces (phase fronts) in the far field, and label the relevant features, both for $k_z < k$ and $k_z > k$. Is the real part of the Poynting's vector normal to the equi-phase surfaces?
(d) Find the electric field due to the same line current source placed on the axis of a cylindrical, perfectly conducting waveguide with circular cross section of radius a.

Problem P4.4

Consider a conducting wire with zero radius and length $2l$ excited by a constant current source I_0. The current density is

$$\overline{J}(\overline{r}') = \hat{z} I_0 \sin\left(kl - k\left|z'\right|\right) \delta(x')\delta(y')$$

which satisfies the boundary condition of zero current at $z' = \pm l$. For the radiation field, show that the vector current moment $\overline{f}(\theta, \phi)$ is

$$\overline{f}(\theta, \phi) = \hat{z} \frac{2I_0}{k \sin^2\theta} [\cos(kl \cos\theta) - \cos(kl)]$$

Find the electric field $\overline{E}$ in the radiation zone. What is $\overline{E}$ at $\theta = 0$ and π? Plot the radiation field pattern for $2l = \lambda/2$ which gives $kl = \lambda/2$. Plot the radiation field pattern for $2l = 3\lambda/2$ and show that the nulls occur at $\cos\theta = \pm 1/3$. To help understand these null positions, consider three colinear dipoles separated by $\lambda/2$ in distance with the center element π out of phase with the two end dipoles. Show that the far field is the sum $e^{-i\pi\cos\theta} - 1 + e^{i\pi\cos\theta}$ which is equal to zero as $\cos\theta = \pm 1/3$.

Problem P4.5

Show that the asymptotic behavior of the radiation resistance R of a thin, linear dipole antenna, for large l/λ, is

$$R \approx 60\{f(l) + (1 + \frac{1}{2}\cos 2kl) \ln kl\}$$

where

$$f(l) = \gamma + \ln 2 + \frac{\gamma}{2}\cos 2kl - \frac{\pi}{4}\sin 2kl$$

Problem P4.6

Find approximate values for the order of modes (that is, u) in the case of a spherical antenna as a limiting case of the biconical antenna when the cone angle θ_0 approaches $\pi/2$.

Problem P4.7

Consider the biconical antenna shown in Figure P4.7, where the conical boundaries are given by $\theta = \theta_0$ and $\theta = \pi - \theta_1$ and where the antenna region is filled with a dielectric of relative permittivity ϵ.

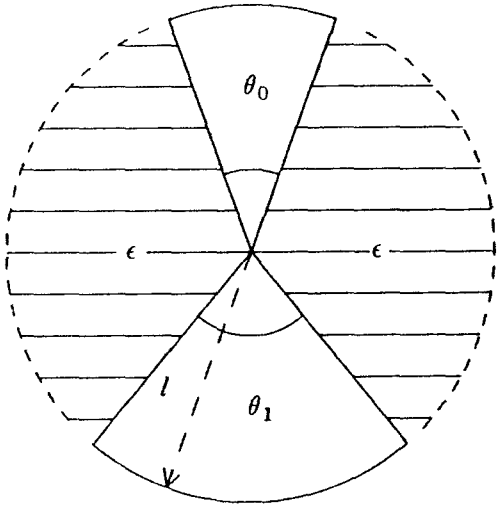


Figure P4.7

(a) Write the boundary conditions for this biconical antenna.
(b) Find the characteristic impedance and prove that the characteristic impedance of a single cone of angle θ_0 over a perfectly conducting ground plane is half that of a biconical antenna with $\theta_0 = \theta_1$.
(c) Find the capacitance and inductance per unit radial length of this biconical antenna.
(d) Find the approximate values of the orders of modes (that is, u) that can be excited in the antenna region for small cone angles θ_0 and θ_1. Prove that these are given by

$$u \simeq n + \Delta$$

where

$$\Delta \simeq -\frac{1}{2} \frac{\ln(\sin\frac{\theta_0}{2} \sin\frac{\theta_1}{2})}{\ln(\sin\frac{\theta_0}{2}) \ln(\sin\frac{\theta_1}{2})}$$

and n is an odd integer.

Problem P4.8

The radiation pattern of the array may also be obtained by using the technique of pattern multiplication. To illustrate the pattern multiplication method, consider four dipoles pointing in the $\hat{z}$ direction, spaced $\lambda/2$ apart with equal amplitude and phase [Fig. P4.8a].

Figure P4.8a

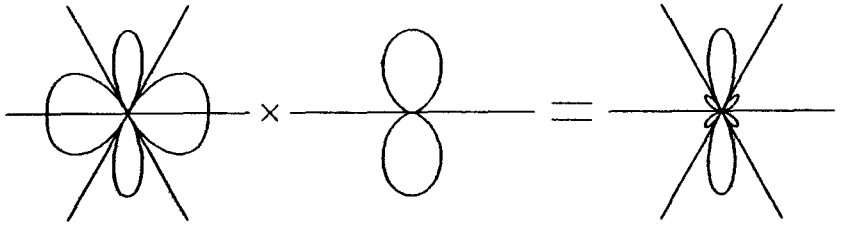


Figure P4.8b

We treat the two elements on the positive and negative x axes as single units, each denoted by the symbol $\odot$ [Fig. P4.8a]. The original array is the convolution (denoted by $\otimes$ in Figure P4.8a) of the unit $\odot$ and the group consisting of two dipoles separated by distance λ. The resultant radiation pattern is the product of the unit pattern and the group pattern shown in Figure P4.8b.

(a) Consider an eight-element linear array with dipoles pointing in the $\hat{z}$ direction and spaced $\lambda/2$ apart with equal amplitude and phase. Using pattern multiplication, sketch the resultant radiation field pattern in the $x-y$ plane. Locate the positions of the nulls and maxima.

(b) Consider four dipoles of equal amplitude and phase on the $x-y$ plane situated at $(x=\lambda/2, y=\lambda/4)$, $(x=\lambda/2, y=-\lambda/4)$, $(x=-\lambda/2, y=\lambda/4)$, and $(x=-\lambda/2, y=-\lambda/4)$.

(c) Using pattern multiplication, sketch the total radiation field pattern in the $x-y$ plane. Locate the positions of the nulls and maxima.

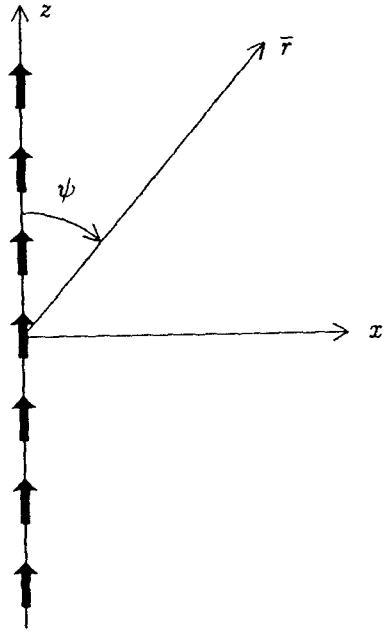


Figure P4.9

Problem P4.9

Consider an equally spaced uniform array of short colinear dipoles with progressive phase shift $\alpha = kd\cos\psi_0$ as shown in Figure P4.9. Show that for $\psi_0 = \pi/2$, the directivity

$$D = \frac{N^2}{W_e}$$

and for $\psi_0 = 0$

$$D = \left|\frac{\sin(Nu/2)}{\sin(u/2)}\sin\psi\right|^2_{\max} \Big/ W_e$$

where

$$W_e = \frac{2N}{3} + 4\sum_{m=1}^{N-1}\frac{N-m}{m^3k^3d^3}\sin(mkd)\cos(mkd\cos\psi_0)$$
$$- 4\sum_{m=0}^{N-1}\frac{N-m}{m^3k^3d^3}\cos(mkd)\cos(mkd\cos\psi_0)$$

Find D as $d \to 0$.

Problem P4.10

Consider a linear broadside array with the array factor

$$\tilde{F}(u) = \sum_{n=0}^{N-1} \sin(nkad) e^{-inu}$$

Notice that the array actually has only $N-1$ elements because the two end elements have zero excitation.

(a) Using z-transformation, show that the magnitude is

$$\left|\tilde{F}(u)\right| = \left| \frac{\cos\dfrac{(N-1)u}{2} \sin\dfrac{\pi}{N-1}}{\cos u - \cos\dfrac{\pi}{N-1}} \right|.$$

(b) For $N = 6$ and $d = \lambda/2$, show that the first sidelobe maximum occurs at $u = \pm 2.44$ for which the sidelobe level is -18 dB.

(c) Show that the directivity is

$$D = \frac{2kd\cot^2(\pi/10)}{\int du \left|\tilde{F}(u)\right|^2}$$

Determine the integration limits for the integral in the denominator. Prove that

$$\begin{aligned} \int du \left|\tilde{F}(u)\right|^2 = 2\sin^2(\pi/5) \Bigg\{ & \frac{2}{3}\sin(3kd)\cos(3\alpha) \\ & + 4\cos(\pi/5)\sin(2kd)\cos(2\alpha) \\ & + 8\cos\frac{\pi}{5}\left[1+\cos\frac{\pi}{5}\right]\sin kd\cos(\alpha) \\ & + \left[6+4\cos\frac{2\pi}{5}\right] kd \Bigg\} \end{aligned}$$

where α is the progressive phase shift between array elements. Show that for $d = \lambda/2$, $D = 3.79$ for all values of α. Calculate the null locations, the first-null beamwidth, and the sidelobe levels

for both broadside and endfire arrays, comparing them with the corresponding (four-element) arrays.

Problem P4.11

Consider a linear antenna with current distribution given by

$$\overline{J} = \hat{z} I_0 \sin\left[k\left(\frac{l}{2} - |z|\right)\right] \delta(x)\delta(y)$$

(a) Show that in the radiation field

$$\overline{E} = \hat{\theta}\left(-\frac{i\eta I_0 e^{ikr}}{2\pi r \sin\theta}\right)\left[\cos\left(\frac{kl}{2}\cos\theta\right) - \cos\left(\frac{kl}{2}\right)\right]$$

$$\overline{H} = \hat{\phi}\frac{1}{\eta}E_\theta$$

(b) Show that the radiation resistance is

$$R_r = \frac{\eta}{2\pi}\int_0^\pi d\theta \frac{[\cos(kl\cos\theta/2) - \cos(kl/2)]^2}{\sin\theta}$$

(c) Plot the radiation resistance as a function of k and show that R_r for a half-wavelength dipole is about 73 Ω.

(d) Calculate the radiation resistance as $kl \to 0$ and compare with that of a Hertzian dipole. Explain the difference.

Problem P4.12

A *turnstile* antenna consists of two Hertzian dipoles positioned at right angles to each other with constant current distributions given by

$$\overline{J}_1 = \hat{x} Il\delta(\bar{r}) \qquad \text{and} \qquad \overline{J}_2 = \hat{y} i Il\delta(\bar{r})$$

respectively.

(a) Show that the electric field produced by this antenna is

$$\overline{E} = -\eta\frac{ikIle^{ikr}}{4\pi r}e^{i\phi}\left\{\hat{r}\left[\frac{i}{kr} + \left(\frac{i}{kr}\right)^2\right]2\cos\theta\right.$$

$$\left. + \hat{\theta}\left[1 + \frac{i}{kr} + \left(\frac{i}{kr}\right)^2\right]\sin\theta - \hat{\phi}\, i\left[1 + \frac{i}{kr} + \left(\frac{i}{kr}\right)^2\right]\right\}$$

(b) Find the total electric field in the far-field ($k_\rho \gg 1$) in the $x-y$ plane with $\theta = \pi/2$. Show that the real space–time dependence of the electric field is of the form $\cos(\omega t + \phi - k_\rho)$. Note that

$$\hat{x} = \hat{r}\cos\phi\sin\theta - \hat{\phi}\sin\phi + \hat{\theta}\cos\phi\cos\theta$$

and

$$\hat{y} = \hat{r}\sin\phi\sin\theta + \hat{\phi}\cos\phi + \hat{\theta}\sin\phi\cos\theta$$

What is the polarization of the radiated wave in the $x-y$ plane?

(c) Find and sketch the radiation *power* pattern in the $x-y$ plane.

(d) Find the total radiated electric field on the z axis. What is the polarization of the radiated wave in the $\hat{z}$ direction?

(e) Calculate the power density radiated in the $+\hat{z}$ direction in the far field and compare it with the radiated power density in the $+\hat{x}$ direction.

Problem P4.13

(a) Find the excitation amplitudes of a five-element broadside Dolph-Chebyshev array with symmetric excitations and with $d = \lambda/2$.

(b) If the sidelobe level $1/R = 0.03$, what is the beamwidth between the first nulls?

(c) For $d = \lambda/3$, what is the beamwidth if $1/R = 0.03$?

(d) Again for $d = \lambda/3$ and $1/R = 0.03$, compare the beamwidth of the Dolph-Chebyshev array with another array generated by the polynomial $P_4(x) = 14.2x^4 - 19.3x^2 + 5.56$ under the Dolph transformation and explain.

(e) For case (c), use Riblet transformation and redesign the Dolph-Chebyshev array.

Problem P4.14

Synthesize a seven-element, equally spaced, broadside array with $d = \lambda/2$, which has -20 dB sidelobe levels. Determine the excitation coefficients. Show that the first-null beamwidth is BW $= 40.2°$ and the directivity is $D = 6.66$ if the array elements are isotropic radiators. Now let $d = \lambda/4$. If the above excitation coefficients are used, show that BW $= 87°$ and $D = 3.3871$.

Problem P4.15

Derive Stirling's formula

$$n! \approx (2\pi)^{1/2} n^{n+1/2} e^{-n}$$

for large n from the defining integral

$$n! = \int_0^\infty dx\, x^n e^{-x}$$

by using the Laplace method.

Problem P4.16

Consider the synthesis of the array pattern [Ma, 1974]

$$G(\xi) = e^{-(\xi-2)^2/4} \qquad -2 \le \xi \le 2$$

with four elements, where $\xi = 2\cos u = 2\cos[kd\cos\psi - \alpha]$.

(a) Let $x = \xi/2$ and choose the zeros of the Legendre polynomial $P_4(x)$ as the sampling points for $g(x) = \exp[-(1-x)^2]$. The zeros are

$$x_0 = -0.861, \qquad x_1 = -0.340, \qquad x_2 = 0.340, \qquad x_3 = 0.861$$

Making use of the Lagrange interpolation formula, show that

$$L(x) = -0.2492x^3 + 0.1588x^2 + 0.7360x + 0.3883$$

which is positive for $-1 \le x \le 1$.

(b) To estimate the maximum error, $\epsilon_{\max}$, note that

$$|\pi(x)|_{\max} = \frac{N!}{(2N-1)(2N-3)\ldots 1} |P_N(x)|_{\max}$$

and that $P_N(x) \le 1$ for $-1 \le x \le 1$. Show that the maximum possible error as estimated is

$$\frac{4!\,12}{7 \cdot 5 \cdot 3 \cdot 1 \cdot 4!} = 0.1143$$

The actual maximum error is $\epsilon(0) = |L(0) - G(0)| \approx 0.0203$.

(c) Show that the power pattern $P(\xi)$ is

$$P(\xi) = -0.0311(\xi - 4.5086)(\xi^2 + 3.2336\xi + 2.7651)$$

which is positive for $-2 \leq \xi \leq 2$. Determine the corresponding array pattern.

Problem P4.17

The method of trigonometric interpolation [Lanczos, 1956] is very useful when the sampling points are equally spaced [Ma, 1974]. With this method it is more convenient to deal with the power patterns as a function of u, $P(u)$. Remember that $\xi = 2\cos u$. The original visible range can be transformed to give the range of u from 0 to π. The sampling positions are obtained by dividing the u range into $N-1$ intervals.

$$u_l = \frac{l\pi}{N-1} \qquad l = 0, 1, \ldots, N-1$$

For every given function of bounded variation $G(u)$, the trigonometric interpolation method provides $P(u)$ which converges unlimitedly to $G(u)$ as the number of sampling points increases.

Consider Problem P4.16 and let $kd = \pi$ and $\alpha = 0$ so that the range of u is $[0, \pi]$ without additional transformations. We take

$$u_0 = 0, \quad u_1 = \pi/3, \quad u_2 = 2\pi/3, \quad u_3 = \pi$$
$$G(u_0) = 1, \quad G(u_1) = 0.779, \quad G(u_2) = 0.105, \quad G(u_3) = 0.018$$

(a) Show that

$$\begin{aligned} P(u) &= \frac{1}{2}a_0 + a_1\cos u + a_2\cos 2u + \frac{1}{2}a_3\cos 3u \\ &= 0.4643 + 0.5520\cos u + 0.0447\cos 2u - 0.06120\cos 3u \end{aligned}$$

(b) Show that, in terms of ξ, the result is

$$P(\xi) = -0.0305(\xi - 4.2893)(\xi^2 + 3.5571\xi + 3.2081)$$

and the power pattern satisfies the realization condition $P(\xi) \geq 0$ for $-2 \leq \xi \leq 2$.

Problem P4.18

Consider the approximation of a given power pattern by the use of Bernstein polynomials. For a real function $g(x)$ defined in the interval $0 \leq x \leq 1$, the Bernstein polynomial approximation of order n of $g(x)$ is [Ma, 1974]

$$B_n^g(x) = \sum_{l=0}^{n} \frac{n!}{l!(n-l)!} g(l/n) x^l (1-x)^{n-l}$$

The convergence of $B_n^g(x)$ to $g(x)$ is also unlimited as n is increased. We see that when $g(x)$ is a constant, $B_n^g(x)$ is equal to the same constant for all n. The most important property of $B_n^g(x)$ in array synthesis is that if $g(x)$ is bounded, $B_n^g(x)$ is bounded with the same upper and lower limits. Thus the resultant approximating polynomial will always satisfy the realization condition while the other interpolation methods provide no such guarantee. For the array pattern as given in Problem P4.16, find the Bernstein polynomial approximation for $n = 3$ and determine the corresponding array pattern.

Problem P4.19

A function $f(\alpha)$ is analytic in a domain D if its derivative $f'(\alpha)$ exists in D. Prove that $f(\alpha) = \alpha^3$ is analytic everywhere while $f(\alpha) = \alpha^*$, where α^* means complex conjugate of α, is non-analytic. Make use of the definition for derivative

$$f'(\alpha) = \lim_{\Delta\alpha \to 0} \frac{(\alpha + \Delta\alpha)^* - \alpha^*}{\Delta\alpha}$$
$$= \lim_{\substack{\Delta_R \to 0 \\ \Delta\alpha_I \to 0}} \frac{\Delta\alpha_R - i\Delta\alpha_I}{\Delta\alpha_R + i\Delta\alpha_I}$$

The derivative $f'(\alpha)$ does not exist if the limit depends on how $\Delta\alpha \to 0$. Find the limit for $f'(\alpha)$ as $\Delta\alpha_I = 0$ and as $\Delta\alpha_R = 0$ and show that $f'(\alpha)$ does not exist because approaching $\alpha = 0$ from the real axis and from the imaginary axis yield different results.

Problem P4.20

(a) Show that if $f(\alpha)$ is analytic in a domain D, then both f_R and f_I satisfy the Cauchy-Riemann equations

$$\frac{\partial f_R}{\partial \alpha_R} = \frac{\partial f_I}{\partial \alpha_I}, \qquad \frac{\partial f_R}{\partial \alpha_I} = -\frac{\partial f_I}{\partial \alpha_R}$$

The converse is true only when these partial derivatives are continuous in D.

(b) Show that if $f(\alpha)$ is analytic in D then both f_R and f_I satisfy the Laplace equation

$$\left\{\frac{\partial^2}{\partial\alpha_R^2}+\frac{\partial^2}{\partial\alpha_I^2}\right\}\begin{bmatrix} f_R(\alpha_R,\alpha_I) \\ f_I(\alpha_R,\alpha_I)\end{bmatrix}=0$$

Problem P4.21

Making use of the Cauchy-Riemann equations and Green's theorem

$$\oint_C d\bar{l}\cdot\overline{A}=\iint_D d\bar{s}\cdot(\nabla\times\overline{A})$$

prove Cauchy's Theorem, which states that if $f(\alpha)$ is analytic in a domain D and on the contour C bounding the domain D, then

$$\oint_C d\alpha\, f(\alpha)=0$$

Problem P4.22

Prove Jordan's Lemma which states that if

$$\lim_{R\to\infty} R\,f(Re^{i\phi})=0$$

then

$$\lim_{R\to\infty}\int_{C_R} d\alpha\, f(\alpha)=0$$

and if

$$\lim_{R\to\infty} f(Re^{i\phi})=0$$

then

$$\lim_{R\to\infty}\int_{C_R} d\alpha\, f(\alpha)e^{ia\alpha}=0$$

for $a>0$, where C_R is a semicircle of radius R in the upper half of the α plane.

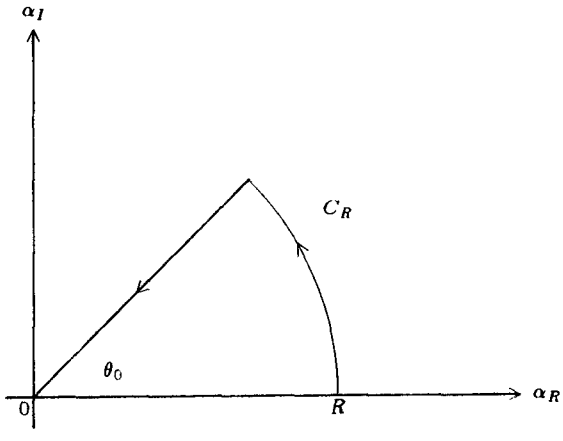


Figure P4.23

Problem P4.23

Using Cauchy's Theorem show that

$$\int_0^\infty \sin \alpha^2 \, d\alpha = \int_0^\infty \cos \alpha^2 \, d\alpha = \sqrt{\pi/8}$$

by writing

$$\int_0^\infty \cos \alpha^2 \, d\alpha + i \int_0^\infty \sin \alpha^2 \, d\alpha = \int_0^\infty e^{i\alpha^2} \, d\alpha$$

and choosing a contour C as illustrated in Figure P4.23. Let $\theta_0 = \pi/4$ and $\alpha = r\, e^{i\pi/4}$ to complete the proof.

Problem P4.24

Show that, by choosing the contour as shown in Figure P4.24

$$I = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} d\alpha \frac{e^{\alpha t}}{\sqrt{\alpha + 1}} = \sqrt{\frac{1}{\pi t}} e^{-t}$$

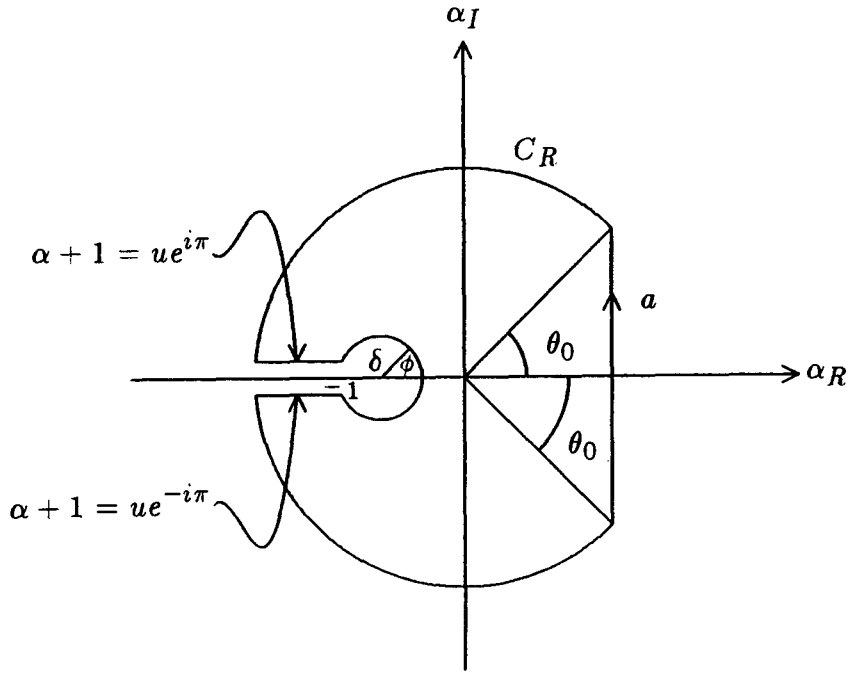


Figure P4.24

Problem P4.25

The Laplace method is useful in evaluating asymptotic values for an integration along the real axis.
Evaluate the integral

$$I(\alpha) = \int_0^\infty dy \frac{y^\alpha e^{-y}}{1+y} \qquad \text{as} \quad \alpha \to \infty$$

Let $y = \alpha x$, and write

$$I(\alpha) = \alpha^{\alpha+1} \int_0^\infty dx \frac{e^{\alpha(-x+\ln x)}}{1+\alpha x}$$

Expand around $x = 1$ and show that

$$I(\alpha) = \sqrt{2\pi\alpha} \frac{\alpha^\alpha}{1+\alpha} e^{-\alpha}$$

Problem P4.26

The stationary-phase method is useful in evaluating asymptotic values for integrals involving rapidly oscillating integrands. Consider the leading behavior for the integral defining the Bessel function

$$J_n(\rho) = Re\left\{\frac{1}{\pi}\int_0^{\pi} e^{i\rho\sin x}e^{-inx}\,dx\right\} \qquad \text{as} \quad \rho \to \infty$$

As $\rho \to \infty$, the term $\exp(i\rho\sin x)$ oscillates very rapidly and tends to cancel out all contributions except near the stationary-phase point $x = \pi/2$ where the oscillation is slowest. The stationary-phase method claims that the largest contribution to the integral comes from near this stationary-phase point $x = \pi/2$. Expanding the function $\sin x$ in the exponential around $\pi/2 - \epsilon \le x \le \pi/2 + \epsilon$ with a very small ϵ, show that

$$\begin{aligned} J_n(\rho) &\simeq Re\left\{\frac{1}{\pi}e^{i(\rho - n\pi/2)}\int_{\pi/2-\epsilon}^{\pi/2+\epsilon} e^{-i(\rho/2)(x-\pi/2)^2}\,dx\right\} \\ &\simeq Re\left\{e^{i(\rho-n\pi/2)}\sqrt{\frac{2}{i\pi\rho}}\right\} = \sqrt{2/\pi\rho}\,\cos(\rho - n\pi/2 - \pi/4) \end{aligned}$$

Problem P4.27

A cylindrical metallic waveguide with a gap at $z = 0$ is excited by a voltage source so that $E_z(a, \phi, z) = V\delta(z)$, where a is the radius of the waveguide. Write the field inside the waveguide as

$$E_z = \int_{-\infty}^{\infty} dk_z g(k_z) J_0(k_\rho \rho) e^{ik_z z}.$$

Show that $g(k_z) = V/[2\pi J_0(k_\rho a)]$. Determine E_z by contour integration.

Problem P4.28

For a linear, temporally dispersive medium with conductivity σ, we have

$$\epsilon(\omega) = \epsilon_0\left[1 + \frac{i\sigma}{\omega\epsilon_0} + \int_0^{\infty} d\tau\, \xi_e(\tau)e^{i\omega\tau}\right]$$

(a) Integrate $\oint_C d\alpha\,[\epsilon(\alpha) - \epsilon_\infty]/(\alpha - \omega)$ over a semicircle of infinite radius with the straight side along the real axis but indented around

the points $\alpha = \omega$ and $\alpha = 0$. Show that the Kramers-Krönig's relation is

$$\epsilon_R(\omega) - \epsilon_\infty = \frac{1}{\pi}\mathrm{PV}\int_{-\infty}^{\infty} d\alpha \frac{\epsilon_I(\alpha)}{\alpha - \omega}$$

$$\epsilon_I(\omega) = -\frac{1}{\pi}\mathrm{PV}\int_{-\infty}^{\infty} a\alpha \frac{\epsilon_R(\alpha) - \epsilon_\infty}{\alpha - \omega} + \frac{\sigma}{\omega}$$

(b) Show that the result in (a) is independent of whether the indentation is made above or below the singularities at $\alpha = \omega$ and $\alpha = 0$.

Problem P4.29

By carrying out the saddle-point evaluation for the H_z component of a VMD on a half-space to a higher order, and combining the leading order branch-point contribution, show that

$$H_z = -i\frac{Ia}{2\pi(k_t^2 - k^2)\rho^2}\left(k_t^3 e^{ik_t\rho} - k^3 e^{ik\rho}\right)$$

Problem P4.30

Consider the integral

$$I(\xi) = \int_\Gamma d\alpha\, F(\alpha) e^{i\xi\cos(\alpha - \alpha_0)}$$

Assuming that both ξ and α_0 are real and that $k = k_R + ik_I$ has a small imaginary part k_I. Show that the steepest descent paths are given by Figure P4.30.

Problem P4.31

Using the following steps, prove the Sommerfeld identity

$$\frac{e^{ik_0 r}}{r} = \frac{i}{2}\int_{-\infty}^{\infty} dk_\rho \frac{k_\rho}{k_z} H_0^{(1)}(k_\rho \rho) e^{ik_z|z|}$$

where $k_z = \sqrt{{k_0}^2 - {k_\rho}^2}$.

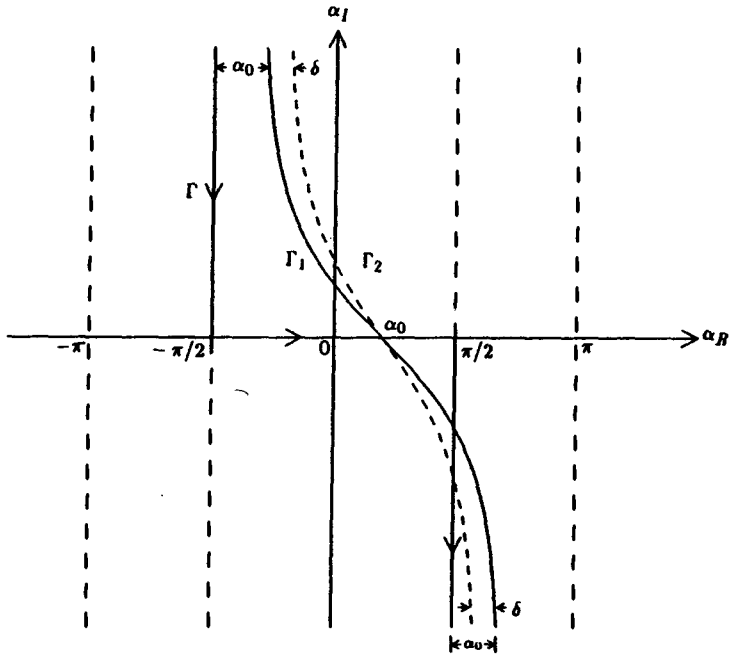


Figure P4.30

(a) Derive an integral representation for $\dfrac{e^{ik_0 r}}{r}$ in terms of its Fourier transform,

$$\frac{e^{ik_0 r}}{r} = \frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} dk_x dk_y dk_z \; g(k_x, k_y, k_z) e^{ik_x x + ik_y y + ik_z z}$$

Assuming that $k_0 = k_0' + ik_0''$, show that

$$g(k_x, k_y, k_z) = \frac{-4\pi}{k_0{}^2 - k_x{}^2 - k_y{}^2 - k_z{}^2}$$

and using contour integration, show that

$$\frac{e^{ik_0 r}}{r} = \frac{i}{2\pi} \iint_{-\infty}^{\infty} dk_x dk_y \, \frac{1}{k_z} \, e^{ik_x x + ik_y y + ik_z |z|}$$

where $k_z = \sqrt{k_0^2 - k_x^2 - k_y^2}$. Note that the above is a statement that a spherical wave can be expanded into a superposition of plane waves.

(b) Transforming the above into cylindrical coordinates and using the identities

$$J_0(k_\rho \rho) = \frac{1}{2}[H_0^{(1)}(k_\rho \rho) + H_0^{(2)}(k_\rho \rho)]$$
$$-H_0^{(1)}(x) = H_0^{(2)}(e^{-i\pi}x)$$

prove the Sommerfeld identity.

Problem P4.32

The two-dimensional and three-dimensional scalar Green's functions as given in the text satisfy the following equations respectively:

Two-dimensional Green's function: $G(\rho) = \frac{i}{4}H_0^{(1)}(k\rho)$

$$\left[\frac{1}{\rho}\frac{d}{d\rho}\left(\rho\frac{d}{d\rho}\right) + k^2\right]G(\rho) = -\frac{\delta(\rho)}{2\pi\rho}$$

Three-dimensional Green's function: $g(\bar{r}) = \frac{e^{ikr}}{4\pi r}$

$$(\nabla^2 + k^2)g(\bar{r}) = -\delta(\bar{r})$$

There are various ways of finding out the relationships between them. First consider the asymptotic behavior of the integral

$$I(r) = \int_{-\infty}^{\infty} dk_\rho\, A(k_\rho)\frac{k_\rho}{k_z}H_0^{(1)}(k_\rho \rho)e^{ik_z z}$$

in the limit when $r \to \infty$. Noting that $k^2 = k_\rho{}^2 + k_z{}^2$ and $r^2 = \rho^2 + z^2$, the following transformations are useful:

$$k_\rho = k\sin\theta, \qquad \rho = r\sin\theta_0$$
$$k_z = k\cos\theta, \qquad z = r\cos\theta_0$$

(a) Show that when calculated to the second order in $1/r$, with the saddle-point method, the asymptotic form for $I(r)$ reads

$$I(r) = \frac{2}{i}\frac{e^{ikr}}{r}\left\{A(\theta_0) - \frac{i}{2kr}\left[A''(\theta_0) + A'(\theta_0)\cot\theta_0\right]\right\}$$

(b) In the case when $A(\theta)$ is a constant, deduce the useful identity in the previous problem,

$$\frac{e^{ikr}}{r} = \frac{i}{2}\int_{-\infty}^{\infty} dk_\rho \frac{k_\rho}{k_z} H_0^{(1)}(k_\rho \rho) e^{ik_z z}$$

(c) Alternatively, consider solving the two differential equations governing the two- and three-dimensional Green's functions by Fourier transforming with respect to z. Show that

$$\frac{e^{ikr}}{r} = \frac{i}{2}\int_{-\infty}^{\infty} dk_z H_0^{(1)}(k_\rho \rho) e^{ik_z z}$$

and

$$H_0^{(1)}(k\rho) = \frac{1}{i\pi}\int_{-\infty}^{\infty} \frac{e^{ikr}}{r} dz$$

By deforming the contour in the k_z plane, and then changing the integration variable to k_ρ for the first integral recover the identity in (b).

Problem P4.33

In dealing with Sommerfeld-type integrals, it is often convenient to use an angular variable α such that

$$k_\rho = k \sin \alpha$$

Let k be real and positive, and

$$x = k_{\rho R}/k = \sin \alpha_R \cosh \alpha_I$$
$$y = k_{\rho I}/k = \cos \alpha_R \sinh \alpha_I$$

(a) Show that the vertical lines (α_R = constant) in the complex α plane are mapped to a family of confocal hyperbolas in the complex k_ρ plane with the foci at $x = \pm 1$ and $y = 0$.

(b) Show that the horizontal lines (α_I = const) in the α plane are mapped to a family of confocal ellipses in the k_ρ plane with the foci at $x = \pm 1$ and $y = 0$.

(c) Show that the ellipses and the hyperbolas intersect each other perpendicularly and that the image of the set $\{-\pi/2 \leq \alpha_R \leq \pi/2, -\infty \leq \alpha_I \leq \infty\}$ is the entire k_ρ plane.

(d) With elliptic coordinates defined as

$$x = h\cosh\xi\cos\eta$$
$$y = h\sinh\xi\cos\eta$$

Find the relations between (ξ,η) and (α_R,α_I).

Problem P4.34

Using the Sommerfeld identity, show that the exact solution for H_z on the surface due to a VMD on a half-space is

$$H_z = -\frac{Ia}{2\pi(k_t^2-k^2)}\left\{\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\frac{k_t^2}{\rho}e^{ik_t\rho}-\frac{k^2}{\rho}e^{ik\rho}\right)\right.$$
$$\left.+\frac{3}{\rho}\left(\frac{\partial^2}{\partial\rho^2}-\frac{1}{\rho}\frac{\partial}{\partial\rho}\right)\left(\frac{1}{\rho^2}e^{ik_t\rho}-\frac{1}{\rho^2}e^{ik\rho}\right)\right\}$$

This result is due to Van der Pol [Baños, 1966].

Problem P4.35

Evaluate the H_z field component of a vertical magnetic dipole (VMD) by using the transformation $k_\rho = k_1\sin\alpha$ instead of $k_\rho = k\sin\alpha$ as shown in the text. Show that the results are identical to those in the text.

Problem P4.36

The modified saddle-point method deals with poles near the saddle point. Assume that after transforming to the s plane, the integral

$$I = \int_{-\infty}^{\infty} ds\, F(s)e^{-s^2}$$

has a pair of poles at $s = \pm s_p$.

(a) Let

$$F(s) = F_s(s) + \frac{2s_pC}{s^2-s_p^2}$$

Find C and show that the integral for $F_s(s)$ can be evaluated with the regular saddle-point method.

(b) Show that

$$I_p = \int_{-\infty}^{\infty} ds \, \frac{2s_p C}{s^2 - s_p^2} \, e^{-s^2} = i2\pi C e^{-s_p^2} \left[1 - \text{erf}(-is_p)\right]$$

where the error function

$$\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z ds \, e^{-s^2}$$

Problem P4.37

Cosider the field near the interface for a vertical electric dipole (VED) placed on the surface of a half-space medium with permittivity ϵ_1 and permeability μ. In region 0, the permittivity is ϵ and the permeability is μ. The vertical electric field component E_z is

$$E_z = -\frac{Il}{8\pi\omega\epsilon} \int dk_\rho \, \frac{k_\rho^3}{k_z} \left(1 + R_{01}^{\text{TM}}\right) H_0^{(1)}(k_\rho \rho) e^{ik_z z}$$

where

$$1 + R_{01}^{\text{TM}} = \frac{2\epsilon_1 k_z}{\epsilon_1 k_z + \epsilon k_{1z}} = \frac{2\epsilon_1 \sqrt{k^2 - k_\rho^2}}{\epsilon_1 \sqrt{k^2 - k_\rho^2} + \epsilon \sqrt{k_1^2 - k_\rho^2}}$$

$k^2 = \omega^2 \mu\epsilon$ and $k_1^2 = \omega^2 \mu\epsilon_1$. Contrary to the case of a vertical magnetic dipole on the half-space medium where no poles exist in the integrand, we now have poles for the VED case.

(a) Show that two simple poles occur at

$$k_p = \pm \frac{k k_1}{\sqrt{k^2 + k_1^2}}$$

These are known as the Sommerfeld poles. When k and k_1 both have non-negative real and imaginary parts, the pole with the plus sign is in the first quadrant of the k_ρ plane and the one with the minus sign is in the third quadrant.

(b) The poles appear on the Riemann sheet for which k_z and k_{1z} are opposite in sign at $k_\rho = k_p$. At the pole location show that

$$k_z = \pm \sqrt{k^2 - k_p^2} = \pm \frac{k_p k}{k_1}$$

$$k_{1z} = \pm\sqrt{k_1^2 - k_p^2} = \pm\frac{k_p k_1}{k}$$

(c) Consider the Sommerfeld pole in the first quadrant. Let both k and k_1 have positive imaginary parts, such that

$$k = |k|\, e^{i\delta}$$
$$k_1 = |k_1|\, e^{i\delta_1}$$

and both δ and δ_1 are positive corresponding to lossy media. Show that the magnitude and phase for k_p are

$$k_p = \frac{|k|\, e^{i\delta}}{\left[1 + A^2 e^{i2(\delta-\delta_1)}\right]^{1/2}}$$
$$= |k_p| \exp\left\{i\delta - \frac{i}{2}\tan^{-1}\frac{A^2 \sin 2(\delta - \delta_1)}{1 + A^2\cos 2(\delta - \delta_1)}\right\}$$

where $A = |k|/|k_1|$ and $|k_p| = |k|\,[1 + 2A^2\cos 2(\delta - \delta_1) + A^4]^{-1/4}$. Show that

$$k_z(k_p) = \pm\frac{|k_p|\,|k|}{|k_1|}\exp\left\{i2\delta - i\delta_1 - \frac{i}{2}\tan^{-1}\frac{A^2\sin 2(\delta-\delta_1)}{1 + A^2\cos 2(\delta-\delta_1)}\right\}$$
$$k_{1z}(k_p) = \pm\frac{|k_p|\,|k_1|}{|k|}\exp\left\{i\delta - \frac{i}{2}\tan^{-1}\frac{A^2\sin 2(\delta-\delta_1)}{1 + A^2\cos 2(\delta-\delta_1)}\right\}$$

(d) When region 0 is free space and medium 1 is slightly lossy, let $\delta = 0$, $\delta_1 \ll 1$, and $A^2 \ll 1$, and show that

$$k_z(k_p) = \pm\frac{|k_p|\,|k|}{|k_1|}\exp\left\{-i\frac{\delta_1}{1 + A^2}\right\}$$
$$k_{1z}(k_p) = \pm\frac{|k_p|\,|k_1|}{|k|}\exp\left\{i\delta_1\frac{1 + 2A^2}{1 + A^2}\right\}$$

(e) When region 0 is free space and medium 1 has a large conductivity such that $k_1 \approx (i\omega\mu\sigma)^{1/2}$, let $\delta = 0$, $\delta_1 = \pi/4$, and $A^2 \ll 1$, and show that

$$k_z(k_p) = \pm\frac{|k_p|\,|k|}{|k_1|}\exp\left\{-i\frac{\pi}{4} + \frac{i}{2}\tan^{-1}A^2\right\}$$
$$k_{1z}(k_p) = \pm\frac{|k_p|\,|k_1|}{|k|}\exp\left\{i\frac{\pi}{4} + \frac{i}{2}\tan^{-1}A^2\right\}$$

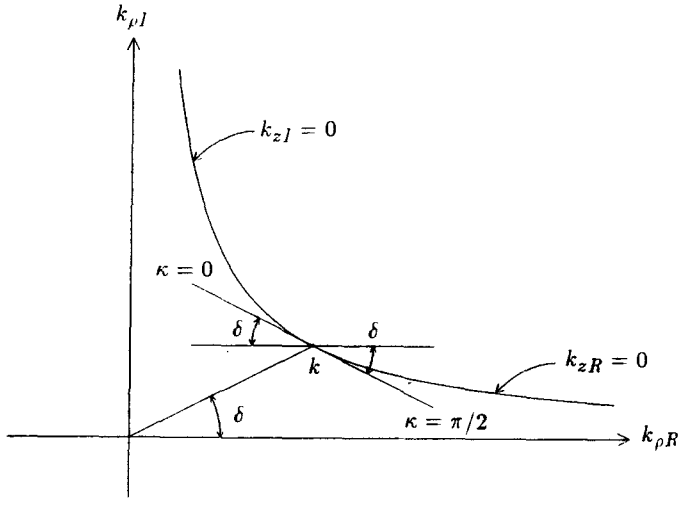


Figure P4.37f

To determine $k_z(k_\rho)$ and $k_{1z}(k_\rho)$ we have to know their allowed phase values on the particular Riemann sheet which is in turn dependent on how the branch cuts are chosen.

(f) To study the various choices of branch cuts originating from k and k_1, let k be complex such that

$$k = k_R + ik_I = |k|\, e^{i\delta}$$

where $k_I \geq 0$ for conductive media. Write

$$\begin{aligned} k_z &= \sqrt{k^2 - k_\rho^2} = |k_z|\, e^{i\kappa} \\ &= \sqrt{(k_R^2 - k_{\rho R}^2) - (k_I^2 - k_{\rho I}^2) + i2(k_R k_I - k_{\rho R} k_{\rho I})} \end{aligned}$$

Show that for the hyperbola on the k_ρ plane determined by

$$k_{\rho R} k_{\rho I} = k_R k_I$$

k_z is either purely real or purely imaginary. The slope of the tangent to the hyperbola at $k_\rho = k_1$ is

$$\frac{dk_{\rho I}}{dk_{\rho R}} = -\frac{k_I}{k_R} = -\tan\delta$$

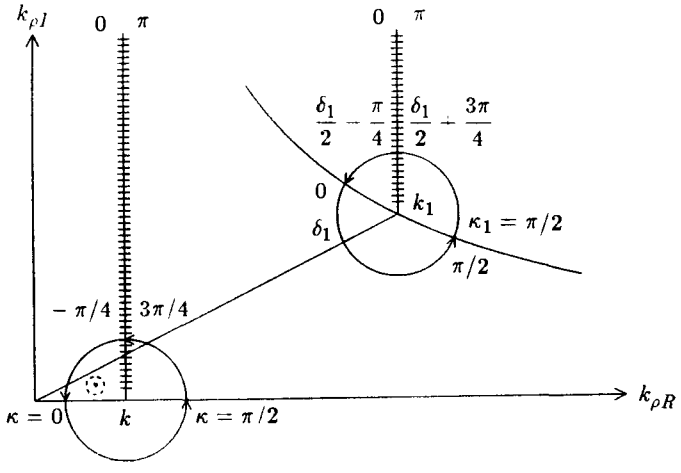


Figure P4.37g

Show that for the part of the hyperbola to the left of k, the curve corresponds to $k_{zI} = 0$. For the part of the hyperbola to the right of k, the curve corresponds to $k_{zR} = 0$ [Fig. P4.37f].

(g) Around the branch point k, let η be a very small number and

$$k_\rho = k + \eta e^{i\beta}$$

Neglecting the η^2 term to obtain

$$k_z = |k_z|\, e^{i\kappa} \approx \sqrt{2\eta\,|k|\, e^{i[\delta+\beta+(2n+1)\pi]}}$$

Thus for $n = 0$ the value of κ increases from 0 to π when β increases from $-\pi - \delta$ to $\pi - \delta$ in the counter-clockwise direction. Consider the branch cuts originating from the branch points k and k_1 by assuming k real and k_1 complex. Let both branch cuts be vertical as shown in Figure P4.37g. The variation of the phase values for k_z and k_{1z}, denoted by κ and κ_1, on the topmost Riemann sheet, are indicated around the two circles centered at k and k_1. Show that both $k_z(k_p)$ and $k_{1z}(k_p)$ take the positive sign in order for their phases κ and κ_1 to be less than $\pi/2$ and larger than $-\pi/4$. Thus the pole is absent.

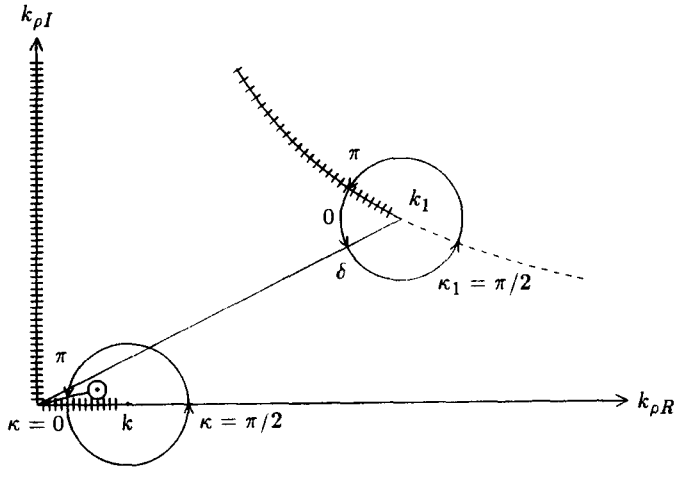


Figure P4.37h

(h) Let both branch cuts follow the constant phase paths as shown in Figure P4.37h. Show that $k_z(k_p)$ takes the minus sign in order for $\pi/2 < \kappa < \pi$ and $k_{1z}(k_p)$ takes the plus sign in order for $0 < \kappa_1 < \pi/2$. Thus the pole is present.

(i) Let the branch cut originating from $k_\rho = k$ follows the $k_{zI} = 0$ path and the branch cut originating from $k_\rho = k_1$ follows the $k_{1zR} = 0$ path as shown in Figure P4.37i. Show that $k_z(k_p)$ takes the minus sign and $k_{1z}(k_p)$ takes the plus sign. Thus the pole is present.

Problem P4.38

Determine the field solutions caused by a horizontal electric dipole placed on the surface of a half-space dielectric medium with permittivity ϵ_t. Plot the radiation patterns inside the dielectric in directions parallel and perpendicular to the dipole axis. Show that the direction for maximum gain is equal to the critical angle in the plane perpendicular to the dipole and is equal to $\sin^{-1}[2/(1 + \epsilon_t/\epsilon)]^{1/2}$ in the plane containing the dipole.

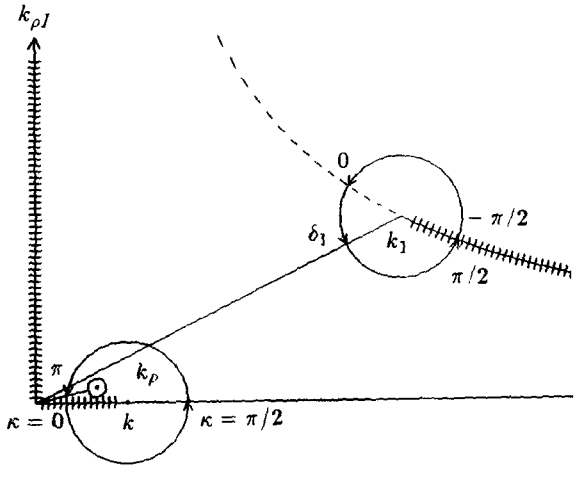


Figure P4.37i

V

THEOREMS OF WAVES AND MEDIA

5.1 Equivalence Principle
- a. Electric and Magnetic Dipole Sources
- b. Image Sources
- c. Electric and Magnetic Current Sheets
- d. Impressed and Induced Current Sheets
- e. Uniqueness Theorem

5.2 Duality and Complementarity

5.3 Mathematical Formulations of Huygens' Principle
- a. Vector Formulations
- b. Scalar Formulation

5.4 Fresnel and Fraunhofer Diffraction

5.5 Reaction and Reciprocity
- a. Reaction
- b. Reciprocity
- c. Modified Reciprocity Theorem

5.6 Stationary Formulas and Rayleigh-Ritz Procedure
- a. Stationary Formula for Resonator Wavenumbers
- b. Stationary Formula for Antenna Impedance
- c. Stationary Formula for Scattering
- d. Method of Moments

5.7 Geometrical Optics Limit

5.8 Paraxial Limit
- a. Gaussian Beam
- b. Gaussian-Hermite Beam Modes
- c. Transmission of Gaussian Beams

5.9 Quasi-Static Limits

5.10 Quantization of Electromagnetic Waves
- a. Uncertainty Principle
- b. Annihilation and Creation Operators
- c. Wave Quantization in Bianisotropic Media

Problems

5.1 Equivalence Principle

When we are interested in a limited region of space, we can replace all uninteresting regions ouside this space by using equivalent sources. We can place equivalent sources in the uninteresting regions, or we can place equivalent current sheets on the boundaries of the region of interest. The equivalent sources are by no means unique, and there are many different ways of constructing them. We need to make sure that all boundary conditions are satisfied and that the original fields and sources in the region of interest are preserved. When two different specifications of sources give the same solution in the region of interest (they certainly will give different solutions outside the region of interest), the two problems are called *equivalent.*

a. Electric and Magnetic Dipole Sources

A small current loop [Fig. 5.1.1a] can be viewed as a magnetic dipole [Fig. 5.1.1b] if the loop is enclosed by a small volume and we are not interested in the interior of the volume. A current loop and a magnetic dipole yields identical results outside the small volume containing the loop and the dipole; only when one penetrates the interior of the sources can one distinguish a current loop from a magnetic dipole. In the interior the magnetic fields of a loop and a magnetic dipole point in opposite directions. Just as electric dipoles constitute the building blocks of electric current sources, the magnetic dipoles constitute the building blocks of magnetic current sources. We denote a magnetic dipole with a double arrow [Fig. 5.1.1c] as opposed to an electric dipole which is denoted with a single arrow [Fig. 5.1.1d].

b. Image Sources

Consider the elementary dipole sources placed in front of a perfect conductor as shown in Figure 5.1.2a. To find solutions in the region of interest, which is the half-space in front of the conductor, we may replace the plane conductor with the images of the dipoles. The image sources thus obtained must satisfy the boundary condition of a zero tangential electric field at the conducting surface. To obtain the image for an electric dipole, we note that the single arrow starts at a negative charge and stops at a positive charge, and that the image of a positive charge is a negative charge and vice versa. To obtain the image for a magnetic dipole, which is representable by a current loop, we note that

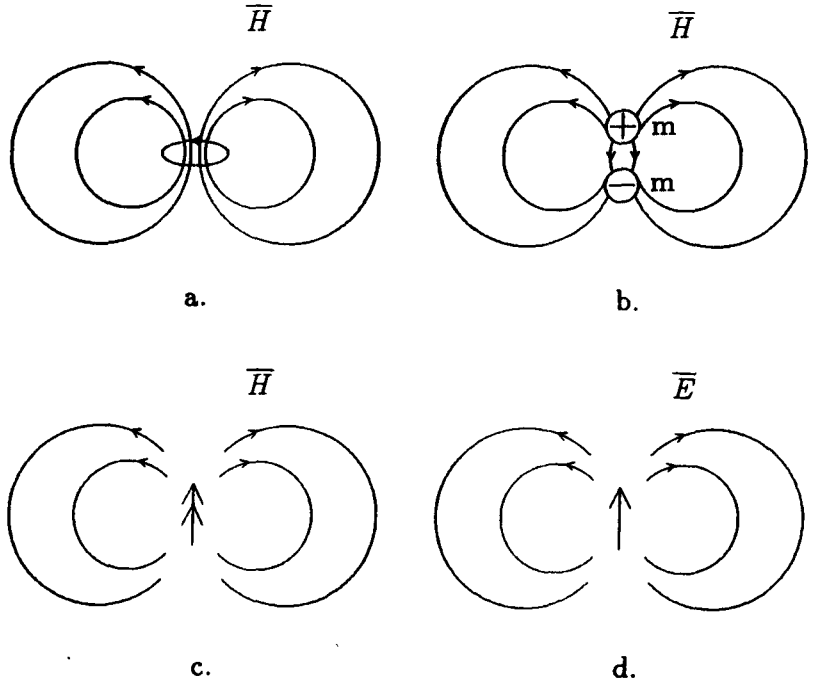


Figure 5.1.1 The equivalence of a. a small current loop, and b. a magnetic dipole. c. A magnetic dipole denoted with a double arrow. d. An electric dipole denoted with a single arrow.

the image of a moving positive charge is a moving negative charge. The images of the four dipoles shown in Figure 5.1.2a is illustrated in Figure 5.1.2b. Notice that with the aid of the image sources, the solution is valid only in the region of interest and does not hold in the image region where the solution should be zero, because it was occupied by the perfect conductor.

As a dual situation we may define a magnetic conductor on the surface of which the tangential magnetic field vanishes. We use wiggly lines to denote a magnetic conductor [Fig. 5.1.3a]. The images of the elementary dipole sources are shown in Figure 5.1.3b.

As a final example of the image method, consider an electric dipole placed between a pair of parallel conducting plates as shown in Figure 5.1.4a. In order to satisfy the boundary conditions at the surfaces of the two plates, we must have multiple image sources as shown in Figure 5.1.4b. The solutions thus obtained are valid only inside the region between the two plates.

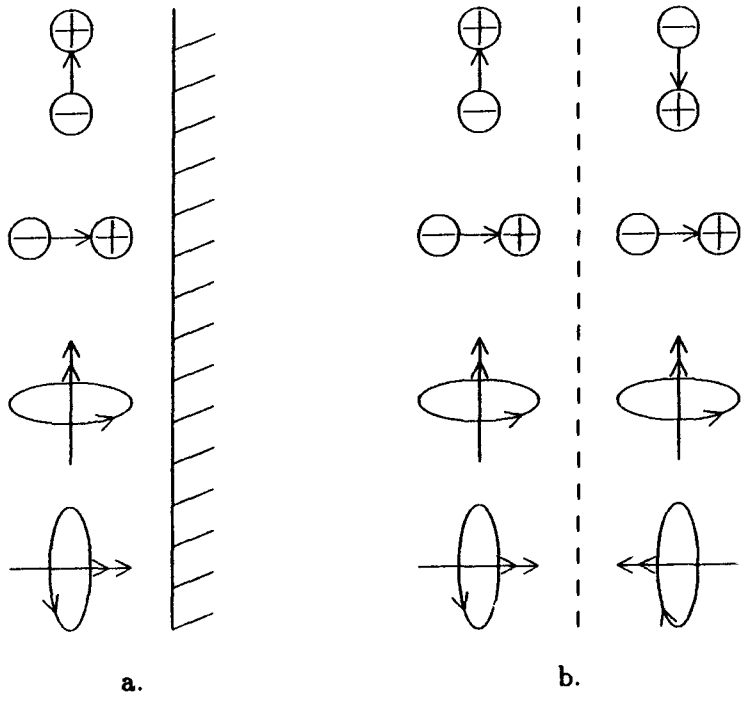


Figure 5.1.2 a. Dipole sources in front of an electric conductor. b. Image sources.

c. Electric and Magnetic Current Sheets

When surface boundaries are replaced by equivalent sources, both electric and magnetic current sheets are required. The electric current sheets $\overline{J}_s$ are produced by discontinuities in tangential magnetic field components across the boundary

$$\overline{J}_s = \hat{n} \times \delta\overline{H}$$

where $\hat{n}$ is the surface normal and $\delta\overline{H}$ is the difference between magnetic field components across the boundary. The magnetic surface current sheets are produced by discontinuities in tangential electric field components across the boundary

$$\overline{M}_s = -\hat{n} \times \delta\overline{E}$$

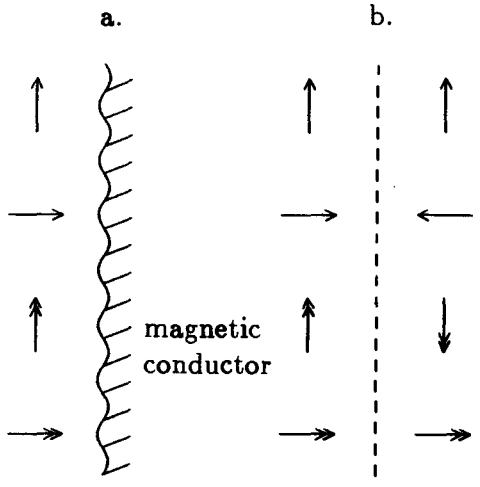


Figure 5.1.3 a. Dipole sources in front of a magnetic conductor. b. Image sources.

Note that, from the definition for $\overline{M}_s$, the circulation of electric fields around $\overline{M}_s$ follows the left-hand rule while the circulation of magnetic fields around $\overline{J}_s$ follows the right-hand rule.

Consider a surface electric current sheet with surface current density $\overline{J}_s = -\hat{x}J_s$ at $z = 0$ [Fig. 5.1.5a]. This current sheet generates plane waves in both the positive and negative $\hat{z}$ directions

$$\overline{E} = \hat{x}\frac{\eta}{2}J_s e^{ikz}, \qquad \overline{H} = \hat{y}\frac{1}{2}J_s e^{ikz} \qquad z > 0$$

$$\overline{E} = \hat{x}\frac{\eta}{2}J_s e^{-ikz}, \qquad \overline{H} = -\hat{y}\frac{1}{2}J_s e^{-ikz} \qquad z < 0$$

We see that the tangential electric fields are continuous at $z = 0$, and that the discontinuity in tangential magnetic fields on both sides equals the strength of the current sheet, $\hat{n} \times \delta\overline{H} = \overline{J}_s$. Thus, all the boundary conditions are satisfied.

As a dual situation, consider a magnetic surface current sheet $\overline{M}_s = -\hat{y}M_s$ at $z = 0$ [Fig. 5.1.5b]. The boundary condition requires that the tangential magnetic fields be continuous across the boundary and that the discontinuity in tangential electric fields be equal to the strength of the current sheet, $-\hat{n} \times \delta\overline{E} = \overline{M}_s$. The solution is as

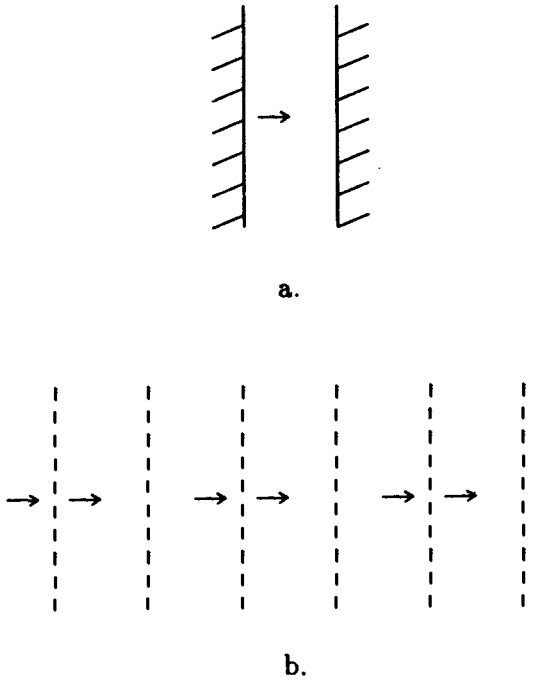


Figure 5.1.4 a. Dipole between two electric conductors. b. Image sources.

follows:

$$\overline{E} = \hat{x}\frac{M_s}{2}e^{ikz}, \qquad \overline{H} = \hat{y}\frac{M_s}{2\eta}e^{ikz} \qquad z > 0$$

$$\overline{E} = -\hat{x}\frac{M_s}{2}e^{-ikz}, \qquad \overline{H} = \hat{y}\frac{M_s}{2\eta}e^{-ikz} \qquad z < 0$$

Plane waves are radiated in both positive and negative $\hat{z}$ directions. Note that, by properly choosing the phase of the current sheets, we can generate plane waves in any direction. For instance, let $\overline{J}_s = \hat{x}J_s e^{ik_y y}$; then plane waves with $\overline{k}$ vectors $\overline{k} = \hat{y}k_y + \hat{z}k_z$ and $\overline{k} = \hat{y}k_y - \hat{z}k_z$ are generated. If k_y is larger than k, the waves are evanescent in the $\hat{z}$ directions. Similar arguments apply to magnetic current sheets.

d. Impressed and Induced Current Sheets

It is important to distinguish between the concepts of *impressed* and *induced* current sheets. On the surface of a material body an induced current sheet is physically carried by charged particles attached

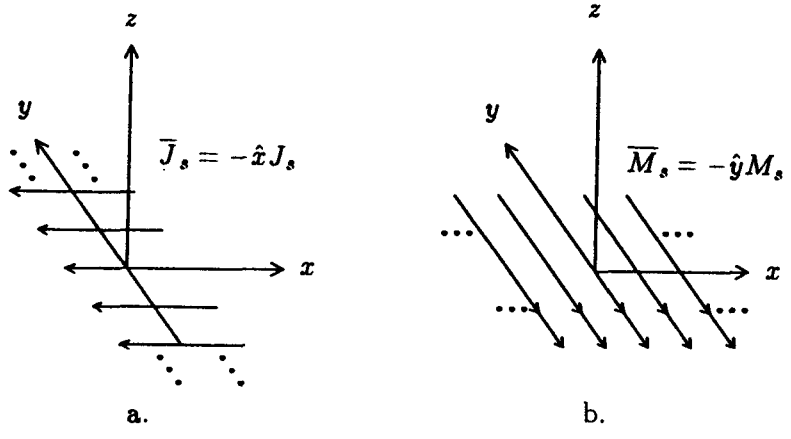


Figure 5.1.5 a. Electric surface currents. b. Magnetic surface currents.

to the surface of the body, whereas an impressed current sheet is carried by external agents. When a layer of charge or current is impressed along the surface of a body, induced surface charge and current sheets are generated at the surface of the body so that the boundary conditions are satisfied.

Consider a plane wave normally incident upon the surface of a perfectly conducting half-space [Fig. 5.1.6a]. Let the electric field be $\overline{E} = \hat{x}E_0e^{ikz}$. An electric current sheet with $\overline{J}_s = \hat{x}2E_0/\eta$ is then induced on the surface of the conductor, and the conductor is replaced by the induced current sheet [Fig. 5.1.6b], which radiates into both the $z > 0$ and $z < 0$ half-space. This induced current generates a reflected wave with $\overline{E} = -\hat{x}E_0e^{-ikz}$, so that at the boundary surface the total electric field is zero. This induced current sheet also generates a plane wave $\overline{E} = -\hat{x}E_0e^{ikz}$ in the region $z > 0$, which combines with the incident wave to produce zero field inside the conductor.

By following the arguments on equivalent sources, it must be appreciated that magnetic sources are useful concepts, although in reality they may not exist. We add an equivalent magnetic source to Faraday's law:

$$\nabla \times \overline{H} = -i\omega\overline{D} + \overline{J} \tag{1}$$

$$-\nabla \times \overline{E} = -i\omega\overline{B} + \overline{M} \tag{2}$$

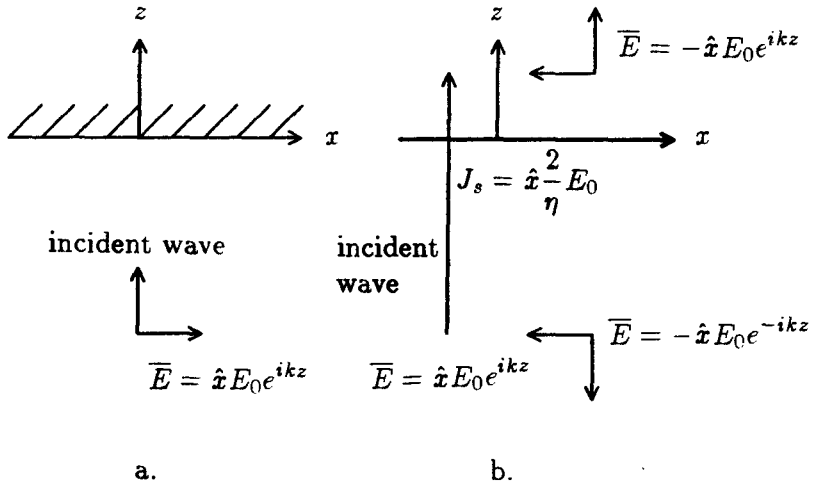


Figure 5.1.6 a. Plane wave incident upon a perfect electric conductor. b. Equivalent electric current sheet placed at $z = 0$.

The added magnetic source is denoted as $\overline{M}$. Similar to the boundary condition for tangential magnetic fields of $\hat{n} \times \overline{H} = \overline{J}_s$, the boundary condition for tangential electric fields becomes $-\hat{n} \times \overline{E} = \overline{M}_s$.

With the use of equivalent current sheet concepts, we now illustrate several equivalent situations for a plane wave propagating in the $\hat{z}$ direction. Let the electric field be $\hat{x}$ directed:

$$\overline{E} = \hat{x}E_0e^{ikz}, \qquad \overline{H} = \hat{y}\frac{1}{\eta}E_0e^{ikz}$$

and let the region of interest be $z > 0$.

Equivalent Problem 1:

Put an electric current sheet with $\overline{J}_s = -\hat{x}E_0/\eta$ and a magnetic current sheet with $\overline{M}_s = -\hat{y}E_0$; then the same field is preserved for $z > 0$. In the region $z < 0$, there is no field.

Equivalent Problem 2:

Put an electric current sheet with $\overline{J}_s = -\hat{x}2E_0/\eta$; then the same field is preserved for $z > 0$. In the negative $\hat{z}$ direction, a plane wave propagates with $\overline{E} = \hat{x}E_0e^{-ikz}$.

Equivalent Problem 3:

Put a magnetic current sheet with $\overline{M}_s = -\hat{y}2E_0$; then the same field is preserved for $z > 0$. In the negative $\hat{z}$ direction, a plane wave propagates with $\overline{H} = \hat{y}(E_0/\eta)e^{-ikz}$.

Equivalent Problem 4:

Replace the region $z < 0$ with a perfect conductor. Place in front of the conductor an electric current sheet with $\overline{J}_s = -\hat{x}E_0/\eta$ and a magnetic current sheet with $\overline{M}_s = -\hat{y}E_0$. The electric current sheet does not generate any field, because an equal and opposite electric current sheet is induced on the surface of the electric conductor and this sheet cancels the impressed $\overline{J}_s$. The magnetic current sheet generates the same field for $z > 0$.

Equivalent Problem 5:

For a dual situation, we place the same electric and magnetic current sheets as in Equivalent Problem 4 in front of a magnetic conductor. According to a similar argument, the magnetic current sheet does not generate a field in the positive $\hat{z}$ direction. The electric current sheet generates the same field for $z > 0$.

From the above discussions on the equivalence principle, we make the following comments: (i) Solutions to the equivalent problems are *not* applicable in the regions of no interest; examples are the $\overline{H}$ field within the small volumes containing a current loop and a magnetic dipole as shown in Figure 5.1.1. (ii) In the case of the image method we turn an unsolved problem of dipoles radiating in the presence of conductors into another unsolved problem of dipole arrays. In the case of placing current sheets on the boundaries of the regions of interest, the current sheet values $-\hat{n} \times \delta\overline{E}$ and $\hat{n} \times \delta\overline{H}$ are unknown until the original problem is completely solved. So what is the usefulness of the equivalence principle? The equivalence principle is useful in at least two aspects. First, it enables us to reformulate a problem as in the case of the image method. Second, and more importantly, it provides us with the means of obtaining approximate solutions by approximating source distributions on the surfaces of regions of interest. However, we must note that equivalent source specifications are by no means unique. For instance, we may use the image sources as shown in Figures 5.1.2–4, or we may place equivalent currents sheets on the surfaces of the conductors. In the following section we discuss the uniqueness theorem which guarantees, at least in the regions of interest, that the solution is unique.

e. Uniqueness Theorem

A given physical situation always leads to one and only one physical solution. However, when formulated in mathematical terms, if not properly done, the problem may lead to many acceptable solutions with underprescribed boundary conditions, or it may permit no solutions at all with overprescribed boundary conditions. The uniqueness theorem indicates how a problem should be properly formulated mathematically so that there is one and only one solution. For electromagnetic field problems, it states that when the sources and the tangential electric *or* magnetic fields are prescribed over the whole boundary surface of a given region, then the solution within this region is unique. The uniqueness theorem is thus a most powerful theorem that enables one to find *the* solution via any expedient means. It is the foundation for the equivalence principle, the Huygens' principle, the image theorem, the induction theorem, Babinet's principle, and almost all frequently used methods in electromagnetism.

To prove the uniqueness theorem, we assume that there are two different solutions for a given set of sources. Let the two solutions be denoted by $\overline{E}_1$ and $\overline{H}_1$, and $\overline{E}_2$ and $\overline{H}_2$. Let the differences be $\delta\overline{E}$ and $\delta\overline{H}$,

$$\delta\overline{E} = \overline{E}_1 - \overline{E}_2$$
$$\delta\overline{H} = \overline{H}_1 - \overline{H}_2$$

Since both field solutions satisfy Maxwell's equations with the same sources, their differences satisfy the source-free Maxwell's equations

$$\nabla \times \delta\overline{E} = i\omega\mu\,\delta\overline{H} \tag{3a}$$
$$\nabla \times \delta\overline{H} = -i\omega\epsilon\,\delta\overline{E} \tag{3b}$$

The proof of the theorem hinges on the assumption that the permittivity ϵ and the permeability μ of the medium have a small imaginary part. Assume the medium is slightly lossy, namely μ and ϵ have a small positive imaginary part,

$$\epsilon = \epsilon_R + i\epsilon_I$$
$$\mu = \mu_R + i\mu_I$$

where ϵ_R, ϵ_I, μ_R and μ_I are real. The proof also holds when the imaginary parts are both negative.

Dot-multiply $(3a)$ with $\delta\overline{H}^*$ and $(3b)^*$ with $\delta\overline{E}$. Subtracting, we obtain

$$\nabla \cdot (\delta\overline{E} \times \delta\overline{H}^*) = i\omega\mu \left|\delta\overline{H}\right|^2 - i\omega\epsilon^* \left|\delta\overline{E}\right|^2 \tag{4}$$

The complex conjugate of (4) gives

$$\nabla \cdot (\delta\overline{E}^* \times \delta\overline{H}) = -i\omega\mu^* \left|\delta\overline{H}\right|^2 + i\omega\epsilon \left|\delta\overline{E}\right|^2 \tag{5}$$

Adding (4) and (5) and integrating over the surface S enclosing the region V, we find

$$\begin{aligned} &\oiint_S d\overline{S} \cdot (\delta\overline{E} \times \delta\overline{H}^* + \delta\overline{E}^* \times \delta\overline{H}) \\ &\qquad = -2\omega \iiint_V dV \left(\mu_I \left|\delta\overline{H}\right|^2 + \epsilon_I \left|\delta\overline{E}\right|^2\right) \end{aligned} \tag{6}$$

The right-hand side of (6) is a negative number. It will be zero if and only if $\delta\overline{H}$ and $\delta\overline{E}$ are identically zero in the region V.

The solution will be unique and $\overline{E}_1 = \overline{E}_2$ and $\overline{H}_1 = \overline{H}_2$ when the left-hand side of (6) is zero, which gives

$$\oiint_S d\overline{S} \cdot (\delta\overline{E} \times \delta\overline{H}^* + \delta\overline{E}^* \times \delta\overline{H}) = 0$$

We conclude that the solution will be unique if either $\delta\overline{E}$ *or* $\delta\overline{H}$ is zero on the enclosed surface S. Thus the boundary conditions can be specified in the following manner: (i) tangential electric field over the whole surface of S, or (ii) tangential magnetic field over the whole surface of S, or (iii) tangential electric field over part of the surface S and tangential magnetic field over the rest of S. If both tangential electric and magnetic fields are specified over any part of S, they must be compatible with each other.

5.2 Duality and Complementarity

Corresponding to the electric current $\overline{J}$ in Ampere's law, a magnetic current $-\overline{M}$ can be added to Faraday's law. Maxwell's equations with both the electric and magnetic current terms read:

$$\nabla \times \overline{H} = -i\omega\epsilon\overline{E} + \overline{J} \tag{1}$$

$$\nabla \times \overline{E} = i\omega\mu\overline{H} - \overline{M} \tag{2}$$

The justification of the magnetic current $\overline{M}$ has been carried out with the use of the equivalence principle and is reiterated as follows. First, (1) and (2) govern *macroscopic* electromagnetic fields under time-harmonic excitations. From the macroscopic point of view, a small current loop acts like a magnetic dipole. As long as one is restricted from penetrating the loop, the fields outside the volume bounding the loop are exactly the dual of those due to a small electric dipole. The fields can be solved in exactly the same manner as in the electric case by neglecting $\overline{J}$ and retaining $\overline{M}$. Second, when (1) and (2) are applied to a limited region in space, the bounding surfaces of the regions can be viewed as supporting surface electric currents due to discontinuities in tangential magnetic field and surface magnetic currents due to discontinuities in tangential electric field. In fact, in the above discussion of dipole sources, we are limited to the space outside the volume occupied by the dipoles. If (1) and (2) are assumed to be valid in all space, then the existence of magnetic monopoles is implied.

Equations (1) and (2) are now duals of each other. If we make the following replacements:

$$\begin{aligned} \overline{E} &\to \overline{H} \\ \overline{H} &\to -\overline{E} \\ \mu &\to \epsilon \\ \epsilon &\to \mu \\ \overline{J} &\to \overline{M} \\ \overline{M} &\to -\overline{J} \end{aligned}$$

then (1) becomes (2) and (2) becomes (1). The symbolic replacements can further be quantified by equating numerically the left-hand side and the right-hand side of the arrows. When such equalities are established, however, the dual of free space will become a medium with permittivity $4\pi \times 10^{-7}$ f/m and with permeability 8.854×10^{-12} h/m, an undesirable situation for the study of antenna problems.

The duality principle that is suitable for antenna and radiation problems in free space is established by the following equalities:

$$\overline{E} = \eta \overline{H} \tag{3}$$

$$\overline{H} = -\overline{E}/\eta \tag{4}$$

$$\overline{J} = \overline{M}/\eta \tag{5}$$

$$\overline{M} = -\eta \overline{J} \tag{6}$$

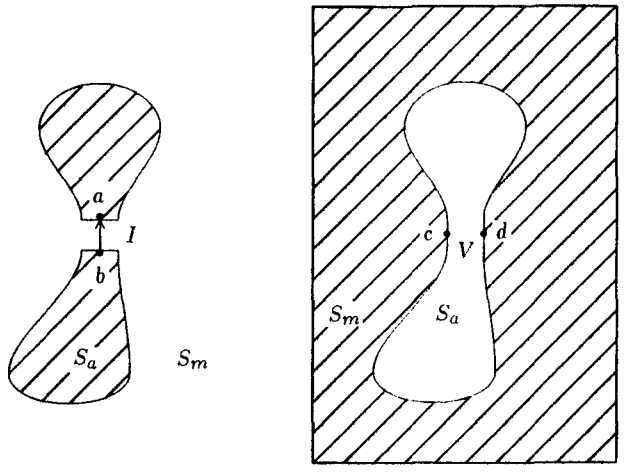


Figure 5.2.1 a. Metal antenna, and b. aperture antenna, which are complementary structures.

where $\eta = (\mu/\epsilon)^{1/2}$. Clearly under such substitutions, (1) becomes (2) and (2) becomes (1). There is no need to replace free space with a different medium. Note that this duality formalism is not applicable to complicated material such as anisotropic and bianisotropic media.

As an example of the application of the duality principle, we shall find a relationship between impedances of complementary metal and aperture antennas. We consider a plane conductor that is cut to make a metal antenna and an aperture antenna complementary to each other as shown in Figure 5.2.1. The input impedance of the metal antenna in Figure 5.2.1a is

$$Z_m = \frac{-\int_b^a d\bar{l} \cdot \overline{E}_m}{\oint_{cd} d\bar{l} \cdot \overline{H}_m} = \frac{-\int_b^a d\bar{l} \cdot \overline{E}_m}{2\int_c^d d\bar{l} \cdot \overline{H}_m} \tag{7}$$

where we have assumed that $\overline{E}_m$ is pointing from a to b and $\overline{H}_m$ from c to d. The second equality follows from the fact that tangential magnetic fields are equal and opposite on both sides of the path cd.

Similarly, the input impedance of the aperture antenna is

$$Z_a = \frac{-\int_c^d d\bar{l} \cdot \overline{E}_a}{\oint_{ab} d\bar{l} \cdot \overline{H}_a} = \frac{\int_c^d d\bar{l} \cdot \overline{E}_a}{2\int_b^a d\bar{l} \cdot \overline{H}_a} \tag{8}$$

where we assume the dual situation with $\overline{H}_a$ pointing from a to b and $\overline{E}_a$ from d to c.

The two impedances are related by duality properties of Maxwell's equations. The boundary-value problem for the metal antenna is to solve

$$\left(\nabla^2 + \omega^2 \mu\epsilon\right) \overline{E}_m = 0 \tag{9}$$

with boundary conditions

$$\begin{aligned} \hat{n} \times \overline{E}_m &= 0 \qquad \text{on} \quad S_a \\ \hat{n} \times \overline{H}_m &= 0 \qquad \text{on} \quad S_m \end{aligned} \tag{10}$$

where

$$\overline{H}_m = \frac{1}{i\omega\mu} \nabla \times \overline{E}_m \tag{11}$$

and $\hat{n}$ is the normal to the plane conductor. Note that the second boundary condition in (10) can be understood by imagining that the surface currents on the metal are composed of elementary current sources with magnetic field having only components perpendicular to S_m.

For the aperture antenna problem, we have to solve

$$\left(\nabla^2 + \omega^2 \mu\epsilon\right) \overline{H}_a = 0 \tag{12}$$

with boundary conditions

$$\begin{aligned} \hat{n} \times \overline{H}_a &= 0 \qquad \text{on} \quad S_a \\ \hat{n} \times \overline{E}_a &= 0 \qquad \text{on} \quad S_m \end{aligned} \tag{13}$$

where

$$\overline{E}_a = -\frac{1}{i\omega\epsilon} \nabla \times H_a \tag{14}$$

The two problems are mathematically dual with the following replacements:

$$\overline{E}_m = \eta \overline{H}_a \tag{15}$$

$$\overline{H}_m = -\frac{1}{\eta} \overline{E}_a \tag{16}$$

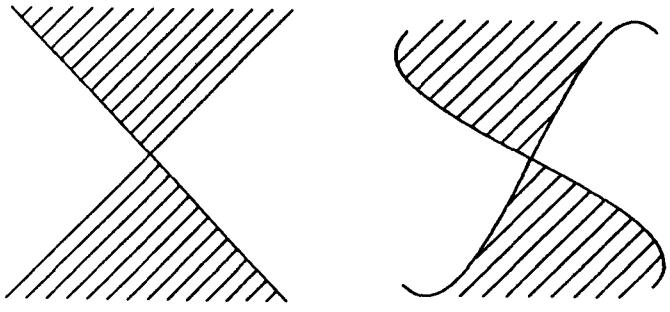


Figure 5.2.2 Broadband structures.

We find from (7) and (8)

$$Z_m Z_a = \left(\frac{-\int_b^a \eta \overline{H}_a \cdot d\bar{l}}{-2\int_c^d \frac{1}{\eta} \overline{E}_a \cdot d\bar{l}} \right) \left(\frac{\int_c^d \overline{E}_a \cdot d\bar{l}}{2\int_b^a \overline{H}_a \cdot d\bar{l}} \right) = \frac{\eta^2}{4} \tag{17}$$

Thus, the product of the input impedances of two planar complementary antennas is one-quarter of the square of the characteristic impedance of the free space.

As an example, consider the structures shown in Figures 5.2.2, where the structure and its complement are identical. It follows that such antennas have input impedance $\eta/2 = 188.5$ ohms. Because the input impedance is independent of frequency, such antennas are ideal broadband antennas. Another self-complementary antenna is shown in Figure 5.2.3. The four edges of the metal are described by the equations $\rho_1 = \rho_0 e^{a\phi}$, $\rho_2 = \rho_0 e^{a(\phi-\pi/2)}$, $\rho_3 = \rho_0 e^{a(\phi-\pi)}$ and $\rho_4 = \rho_0 e^{a(\phi-3\pi/2)}$, where $1/a$ is the rate of expansion of the spiral. The structure is known as a planar equiangular spiral antenna.

Babinet's principle is another example of duality and complementarity that relates the problem of diffraction by planar apertures to the problem of scattering by its complementary structure. Consider an infinite plane conductor with an aperture as shown in Figure 5.2.5a and the complementary structure consisting of the removed plane metal in the formation of the aperture as shown in Figure 5.2.5b. Let there be dual sources on the left-hand sides of Figures 5.2.5a and 5.2.5b. In the absence of the screens, they produce incident fields such that

$$\overline{E}_2^i = \eta \overline{H}_1^i \tag{18a}$$

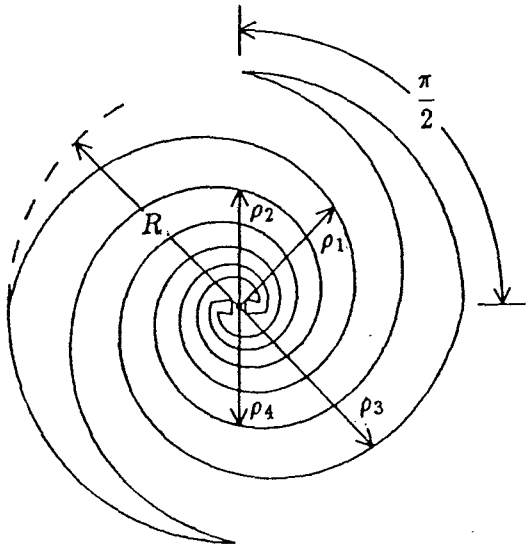


Figure 5.2.3 Planar equiangular spiral antenna.

$$\overline{H}_2^i = -\overline{E}_1^i/\eta \tag{18b}$$

Note that in the presence of the screens the sources are located on the left-hand side.

Now we formulate the problems for the fields on the right-hand sides of the screens. For the problem in Figure 5.2.4a, the fields satisfy Maxwell's equations

$$\nabla \times \overline{E}_1 = i\omega\mu\overline{H}_1 \tag{19a}$$

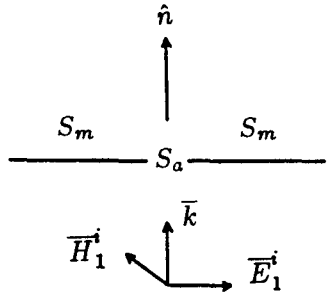
$$\nabla \times \overline{H}_1 = -i\omega\epsilon\overline{E}_1 \tag{19b}$$

subject to the boundary conditions

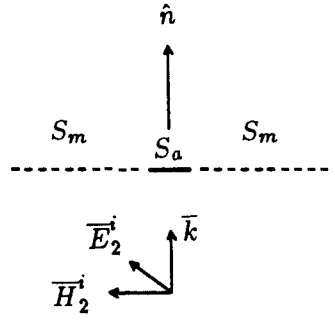
$$\hat{n} \times \overline{E}_1 = 0 \qquad \text{on} \quad S_m \tag{20a}$$

$$\hat{n} \times \overline{H}_1 = \hat{n} \times \overline{H}_1^i \qquad \text{on} \quad S_a \tag{20b}$$

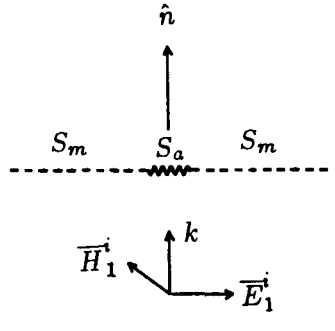
where $\hat{n}$ is the normal to the plane surface. The first boundary condition warrants that tangential electric field vanishes on the conducting surface. The second condition is due to the fact that induced surface currents on the metal surface produce no tangential magnetic field component at the aperture space. The field to the right-hand side of



a. Wave incident upon an electric screen.



b. Complementary problem of a.



c. Dual problem of b.

Figure 5.2.4 Complementarity and duality for Babinet's principle.

the screen is produced by the equivalent current sheet source $\overline{J}_1^i$ on S_a.

For the problem in Figure 5.2.4b, the fields on the right-hand side satisfy Maxwell's equations

$$\nabla \times \overline{E}_2 = i\omega\mu\overline{H}_2 \tag{21a}$$
$$\nabla \times \overline{H}_2 = -i\omega\epsilon\overline{E}_2 \tag{21b}$$

subject to the boundary conditions

$$\hat{n} \times \overline{H}_2 = \hat{n} \times \overline{H}_2^i \qquad \text{on} \quad S_m \tag{22a}$$
$$\hat{n} \times \overline{E}_2 = 0 \qquad \text{on} \quad S_a \tag{22b}$$

The total fields $\overline{E}_2$ and $\overline{H}_2$ are superpositions of two field solutions:

1. The incident fields $\overline{E}_2^i$ and $\overline{H}_2^i$ in the absence of the screens. They satisfy source-free Maxwell's equations to the right half-space.
2. The scattered fields produced by induced currents on the screens. We denote them by $\overline{E}_2^s$ and $\overline{H}_2^s$.

In terms of the scattered field $\overline{E}_2^s = \overline{E}_2 - \overline{E}_2^i$ and $\overline{H}_2^s = \overline{H}_2 - \overline{H}_2^i$ they satisfy

$$\nabla \times \overline{E}_2^s = i\omega\mu\overline{H}_2^s \tag{23a}$$
$$\nabla \times \overline{H}_2^s = -i\omega\epsilon\overline{E}_2^s \tag{23b}$$

subject to the boundary conditions

$$\hat{n} \times \overline{H}_2^s = 0 \qquad \text{on} \quad S_m \tag{24a}$$
$$\hat{n} \times \overline{E}_2^s = -\hat{n} \times \overline{E}_2^i = -\hat{n} \times \eta\overline{H}_1^i \qquad \text{on} \quad S_a \tag{24b}$$

Comparing (23) with (19) and (24) with (20), we see that the two problems are mathematically dual with the following substitutions

$$\overline{H}_2^s = \overline{E}_1/\eta \tag{25a}$$
$$\overline{E}_2^s = -\eta\overline{H}_1 \tag{25b}$$

In terms of total fields, we find

$$\overline{E}_2 = \overline{E}_2^s + \overline{E}_2^i = \eta(-\overline{H}_1 + \overline{H}_1^i) \tag{26a}$$
$$\overline{H}_2 = \overline{H}_2^s + \overline{H}_2^i = \frac{1}{\eta}(\overline{E}_1 - \overline{E}_1^i) \tag{26b}$$

This is referred to as Babinet's principle. Note that in the special case of no metallic screens for Figure 5.2.4a, $\overline{E}_1 = \overline{E}_1^i$ and $\overline{H}_1 = \overline{H}_1^i$. The complementary case is a complete metallic screen and the result is $\overline{E}_2 = \overline{H}_2 = 0$. Consider the other extreme when Figure 5.2.4a is completely metallic with no apertures, $\overline{E}_1 = \overline{H}_1 = 0$. Then the result for Figure 5.2.4b becomes $\overline{E}_2 = \eta \overline{H}_1^i$ and $\overline{H}_2 = -\overline{E}_1^i/\eta$, which are just the fields generated by the dual sources.

To examine further the implications of Babinet's principle, consider the dual problem of Figure 5.2.4b as illustrated in Figure 5.2.4c. The metallic aperture S_a is now replaced by a magnetic conductor and the sources are the dual of Figure 5.2.4c and therefore identical to those of Figure 5.2.4b. the boundary-value problem for the right-hand side of the magnetic conductor becomes

$$\nabla \times \overline{E}_d = i\omega\mu \overline{H}_d \tag{27a}$$
$$\nabla \times \overline{H}_d = -i\omega\epsilon \overline{E}_d \tag{27b}$$

subject to the boundary conditions

$$\hat{n} \times \overline{E}_d = \hat{n} \times \overline{E}_1^i \qquad \text{on} \quad S_m \tag{28a}$$
$$\hat{n} \times \overline{H}_d = 0 \qquad \text{on} \quad S_a \tag{28b}$$

where $\overline{E}_d$ and $\overline{H}_d$ denote fields of Figure 5.2.8. From (19)–(20) and (27)–(28) we see that the sums of the fields $\overline{E} = \overline{E}_1 + \overline{E}_d$ and $\overline{H} = \overline{H}_1 + \overline{H}_d$ satisfy the following boundary-value problem

$$\nabla \times (\overline{E}_1 + \overline{E}_d) = i\omega\mu(\overline{H}_1 + \overline{H}_d) \tag{29a}$$
$$\nabla \times (\overline{H}_1 + \overline{H}_d) = -i\omega\epsilon(\overline{E}_1 + \overline{E}_d) \tag{29b}$$

with the boundary conditions

$$\hat{n} \times (\overline{E}_1 + \overline{E}_d) = \hat{n} \times \overline{E}_1^i \qquad \text{on} \quad S_m \tag{30a}$$
$$\hat{n} \times (\overline{H}_1 + \overline{H}_d) = \hat{n} \times \overline{H}_1^i \qquad \text{on} \quad S_a \tag{30b}$$

Thus the tangential fields on S_m and S_a are identical to those of $\overline{E}_1^i$ and $\overline{H}_1^i$ and by the uniqueness theorem, we conclude that

$$\overline{E}_1 + \overline{E}_d = \overline{E}_1^i \tag{31a}$$
$$\overline{H}_1 + \overline{H}_d = \overline{H}_1^i \tag{31b}$$

This also follows from Babinet's principle as expressed in (26) with the substitution of $\overline{E}_2 = \eta \overline{H}_d$ and $\overline{H}_2 = -\overline{E}_d/\eta$ as required by the duality principle.

5.3 Mathematical Formulations of Huygens' Principle

Huygens' principle states that the field solution in a region V' is completely determined by the tangential fields specified over the surface S' enclosing V' [Fig. 5.3.1]. Formulated in mathematical terms, Huygens' principle expresses fields at an observation point in terms of fields at the boundary surface. Consider a surface S' enclosing a radiating source. The electric and magnetic fields outside the surface will be shown to be of the following forms:

$$\overline{E}(\overline{r}) = \oiint_{S'} dS' \left\{ i\omega\mu \overline{\overline{G}}(\overline{r},\overline{r}') \cdot [\hat{n} \times \overline{H}(\overline{r}')] + \nabla \times \overline{\overline{G}}(\overline{r},\overline{r}') \cdot [\hat{n} \times \overline{E}(\overline{r}')] \right\} \tag{1}$$

$$\overline{H}(\overline{r}) = \oiint_{S'} dS' \left\{ -i\omega\epsilon \overline{\overline{G}}(\overline{r},\overline{r}') \cdot [\hat{n} \times \overline{E}(\overline{r}')] + \nabla \times \overline{\overline{G}}(\overline{r},\overline{r}') \cdot [\hat{n} \times \overline{H}(\overline{r}')] \right\} \tag{2}$$

where $\hat{n}$ is the outward normal to the surface S'. The dyadic Green's function $\overline{\overline{G}}(\overline{r},\overline{r}')$ is given by

$$\overline{\overline{G}}(\overline{r},\overline{r}') = \left[\overline{\overline{I}} + \frac{1}{k^2} \nabla\nabla \right] g(\overline{r},\overline{r}') \tag{3}$$

The scalar Green's function $g(\overline{r},\overline{r}')$ satisfies the Helmholtz equation

$$(\nabla^2 + k^2) g(\overline{r},\overline{r}') = -\delta(\overline{r} - \overline{r}') \tag{4}$$

For three-dimensional problems, the scalar Green's function $g(\overline{r},\overline{r}')$ for isotropic media written in spherical coordinates is of the form

$$g(\overline{r},\overline{r}') = \frac{e^{ik|\overline{r}-\overline{r}'|}}{4\pi|\overline{r}-\overline{r}'|} \tag{5}$$

For two-dimensional problems, the scalar Green's function for isotropic media written in cylindrical coordinates is of the form

$$g(\overline{\rho},\overline{\rho}') = \frac{i}{4} H_0^{(1)}(k|\overline{\rho}-\overline{\rho}'|) \tag{6}$$

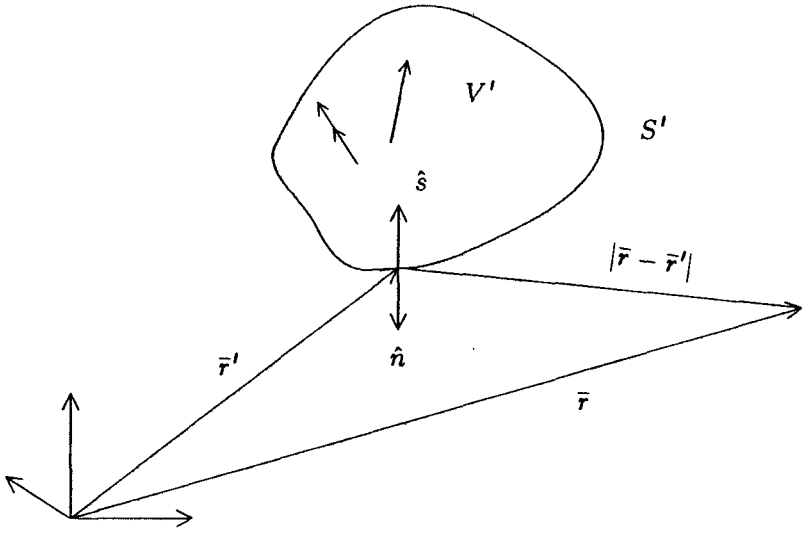


Figure 5.3.1 Volume V' containing radiation sources.

where $H_0^{(1)}$ is the zeroth-order Hankel function of the first kind.

From Maxwell's equations, the governing equation for the electric field $\overline{E}(\bar{r})$ with a given current source $\overline{J}(\bar{r})$ is given by

$$\nabla \times \nabla \times \overline{E}(\bar{r}) - k^2 \overline{E}(\bar{r}) = i\omega\mu \overline{J}(\bar{r}) \tag{7}$$

The solution of $\overline{E}(\bar{r})$ in terms of $\overline{J}(\bar{r})$ can be conveniently expressed in terms of the dyadic Green's function $\overline{\overline{G}}(\bar{r}, \bar{r}')$,

$$\overline{E}(\bar{r}) = i\omega\mu \iiint d^3\bar{r}' \, \overline{\overline{G}}(\bar{r}, \bar{r}') \cdot \overline{J}(\bar{r}') \tag{8}$$

Notice that by means of the three-dimensional delta function $\delta(\bar{r} - \bar{r}')$, we can write

$$\overline{J}(\bar{r}) = \iiint d^3\bar{r}' \, \delta(\bar{r} - \bar{r}') \overline{\overline{I}} \cdot \overline{J}(\bar{r}') \tag{9}$$

where $\overline{\overline{I}}$ is the unit dyad. Substitution of (8) and (9) in (7) gives the following equation governing the dyadic Green's function $\overline{\overline{G}}(\bar{r}, \bar{r}')$,

$$\nabla \times \nabla \times \overline{\overline{G}}(\bar{r}, \bar{r}') - k^2 \overline{\overline{G}}(\bar{r}, \bar{r}') = \overline{\overline{I}} \, \delta(\bar{r} - \bar{r}') \tag{10}$$

The derivation of the dyadic form of Huygens' principle calls for the vector identity

$$\begin{aligned}\overline{E}\cdot[\nabla\times\nabla\times(\overline{\overline{G}}\cdot\overline{a})]-[\nabla\times\nabla\times\overline{E}]\cdot(\overline{\overline{G}}\cdot\overline{a})\\ =-\nabla\cdot\left\{\overline{E}\times\nabla\times\left(\overline{\overline{G}}\cdot\overline{a}\right)+(\nabla\times\overline{E})\times\left(\overline{\overline{G}}\cdot\overline{a}\right)\right\}\end{aligned}$$

where $\overline{a}$ is an arbitrary constant vector. Integrating over a volume V bounded by the closed surface S' and the surface at infinity, we obtain

$$\begin{aligned}\iiint_V d^3\overline{r}\left\{\overline{E}(\overline{r})\cdot[\nabla\times\nabla\times\overline{\overline{G}}(\overline{r},\overline{r}')\cdot\overline{a}]\right.\\ \left.-[\nabla\times\nabla\times\overline{E}(\overline{r})]\cdot\overline{\overline{G}}(\overline{r},\overline{r}')\cdot\overline{a}\right\}\\ =-\oiint_S dS\left\{\hat{s}\cdot[\nabla\times\overline{E}(\overline{r})]\times\overline{\overline{G}}(\overline{r},\overline{r}')\cdot\overline{a}\right.\\ \left.+\hat{s}\cdot\overline{E}(\overline{r})\times\nabla\times\overline{\overline{G}}(\overline{r},\overline{r}')\cdot\overline{a}\right\}\end{aligned}$$

Making use of (7) and (10) in the left-hand side of the above equation and assuming that $\overline{J}(\overline{r})=0$ in volume V, we find

$$\begin{aligned}\overline{E}(\overline{r}')\cdot\overline{a}=-\oiint_S dS\left\{\hat{s}\times[\nabla\times\overline{E}(\overline{r})]\cdot\overline{\overline{G}}(\overline{r},\overline{r}')\cdot\overline{a}\right.\\ \left.+[\hat{s}\times\overline{E}(\overline{r})]\cdot\nabla\times\overline{\overline{G}}(\overline{r},\overline{r}')\cdot\overline{a}\right\}\end{aligned}\tag{11}$$

We now use $\nabla\times\overline{E}(\overline{r})=i\omega\mu\overline{H}(\overline{r})$ and interchange the primed and unprimed variables. Since $\overline{a}$ is an arbitrary constant vector, we can delete it from both sides of the equation. Thus

$$\begin{aligned}\overline{E}(\overline{r})=-\oiint_{S'} dS'\left\{i\omega\mu\overline{\overline{G}}(\overline{r},\overline{r}')\cdot[\hat{s}\times\overline{H}(\overline{r}')]\right.\\ \left.+\nabla\times\overline{\overline{G}}(\overline{r},\overline{r}')\cdot[\hat{s}\times\overline{E}(\overline{r}')]\right\}\end{aligned}\tag{12}$$

In arriving at (12) we also made use of the symmetry relations of the dyadic Green's function:

$$\overline{\overline{G}}(\overline{r},\overline{r}')=\left[\overline{\overline{G}}(\overline{r}',\overline{r})\right]^{\mathrm{T}}\tag{13}$$

and

$$\nabla \times \overline{\overline{G}}(\bar{r}, \bar{r}') = \left[\nabla' \times \overline{\overline{G}}(\bar{r}', \bar{r})\right]^{\mathrm{T}} \tag{14}$$

where superscript T denotes transpose. In view of (3) and (5) or (6), the relation (13) is evidently true. The relation (14) is also seen to be true by noting that

$$\begin{aligned}
[\nabla \times \overline{\overline{G}}(\bar{r}, \bar{r}')]_{il} &= \epsilon_{ijk}\, \partial_j \left(\delta_{kl} + \frac{1}{k^2}\partial_k \partial_l\right) g(\bar{r} - \bar{r}') \\
&= -\epsilon_{ijl}\, \partial'_j\, g(\bar{r} - \bar{r}') \\
&= \epsilon_{ljk}\, \partial'_j\, \delta_{ki}\, g(\bar{r} - \bar{r}') \\
&= [\nabla' \times \overline{\overline{G}}(\bar{r}, \bar{r}')]_{li}
\end{aligned} \tag{15}$$

where we use the relations $\epsilon_{ijk}\partial_j\partial_k = 0$ and $\partial_j g = -\partial'_j g$. It can also be shown that for general dyadic Green's functions satisfying prescribed boundary conditions, the symmetry relations (13) and (14) also hold.

The Huygens' principle is formulated with the identification of the closed surface S comprising a sphere of infinite radius and a surface S' enclosing all sources of radiation [Fig. 5.3.1]. The surface at infinity gives no contribution to the surface integral. This follows from the radiation conditions for the electromagnetic fields and the dyadic Green's functions:

$$\lim_{r\to\infty} r[\overline{H} - \hat{r} \times \overline{E}/\eta] = 0 \tag{16a}$$

$$\lim_{r\to\infty} \bar{r} \cdot \overline{E} = 0 \tag{16b}$$

$$\lim_{r\to\infty} r[\nabla \times \overline{\overline{G}} - ik\hat{r} \times \overline{\overline{G}}] = 0 \tag{16c}$$

Note that $\hat{r} = \hat{s}$ in (12). The term $\hat{s} \times \overline{H}(\bar{r}')$ in (12) becomes $\hat{r} \times (\hat{r} \times \overline{E}/\eta)$ and the integrand becomes $[-ik\hat{r} \times \overline{\overline{G}} + \nabla \times \overline{\overline{G}}] \cdot [\hat{r} \times \overline{E}]$ which vanishes on the surface at infinity. The contribution to $\overline{E}(\bar{r})$ comes solely from the surface S' [Fig. 5.3.1]. Thus, noticing that $\hat{s} = -\hat{n}$, we obtain (1) and (2) which are expressed in terms of the tangential electric and magnetic field components at the boundary surface S'.

a. *Vector Formulations*

Many different vector forms can be obtained for the Huygens' principle. Making use of (3), we obtain from (1) and (2)

$$\overline{E}(\bar{r}) = \iint_{S'} dS' \left\{ i\omega\mu g(\bar{r},\bar{r}')[\hat{n} \times \overline{H}(\bar{r}')] + \frac{i\omega\mu}{k^2} \nabla\nabla g(\bar{r},\bar{r}') \cdot [\hat{n} \times \overline{H}(\bar{r}')] + \nabla g(\bar{r},\bar{r}') \times [\hat{n} \times \overline{E}(\bar{r}')] \right\} \tag{17}$$

$$\overline{H}(\bar{r}) = \iint_{S'} dS' \left\{ -i\omega\epsilon g(\bar{r},\bar{r}')[\hat{n} \times \overline{E}(\bar{r}')] - \frac{i\omega\epsilon}{k^2} \nabla\nabla g(\bar{r},\bar{r}') \cdot [\hat{n} \times \overline{E}(\bar{r}')] + \nabla g(\bar{r},\bar{r}') \times [\hat{n} \times \overline{H}(\bar{r}')] \right\} \tag{18}$$

The above formulas are in terms of the tangential field components on S_1 and the scalar Green's function $g(\bar{r},\bar{r}')$.

The Stratton-Chu formulas can be derived from (11) by noting that the first term inside the surface integral is

$$\hat{s} \times (\nabla \times \overline{E}) \cdot \overline{\overline{G}} \cdot \bar{a} = \hat{s} \times (\nabla \times \overline{E}) \cdot [g\bar{a} + \frac{1}{k^2} \nabla\nabla \cdot g\bar{a}] \tag{19a}$$

$$= \hat{s} \times (\nabla \times \overline{E}) \cdot g\bar{a} + \frac{1}{k^2} \hat{s} \cdot (\nabla \times \overline{E}) \times \nabla\nabla \cdot g\bar{a} \tag{19b}$$

$$= \hat{s} \times (\nabla \times \overline{E}) \cdot g\bar{a} + \frac{1}{k^2} \hat{s} \cdot \{(\nabla \times \nabla \times \overline{E})\nabla \cdot g\bar{a} - \nabla \times [(\nabla \cdot g\bar{a})(\nabla \times \overline{E})]\} \tag{19c}$$

In the above derivations, (19) is obtained by substitution of (3). A vector identity $(\nabla \times \overline{E}) \times \nabla\phi = \phi(\nabla \times \nabla \times \overline{E}) - \nabla \times (\phi\nabla \times \overline{E})$ is used to convert the second term in (19*b*) to the second term in (19*c*). Substituting (19*c*) in (11) we note that the last term in (19*c*) vanishes under the surface integration because by Gauss' theorem it is equivalent to the volume integral of the divergence of a curl. Also note that $\nabla \times \overline{E} = i\omega\mu\overline{H}$ and that for $\overline{J}(\bar{r}) = 0$, $\nabla \times \nabla \times \overline{E}(\bar{r}) = k^2\overline{E}(\bar{r})$. Thus

$$\overline{E}(\bar{r}') = -\iint dS' \left\{ i\omega\mu[\hat{s} \times \overline{H}(\bar{r})]g(\bar{r},\bar{r}') + [\hat{s} \cdot \overline{E}(\bar{r})]\nabla g(\bar{r},\bar{r}') + [\hat{s} \times \overline{E}(\bar{r})] \times \nabla g(\bar{r},\bar{r}') \right\} \tag{20}$$

Interchanging primed and unprimed quantities, making use of the symmetry properties of $g(\bar{r},\bar{r}')$, and applying to the surface S' as shown

in Fig. 5.3.1, we obtain the Stratton-Chu formula

$$\overline{E}(\bar{r}) = \iint_{S'} dS' \left\{ i\omega\mu[\hat{n} \times \overline{H}(\bar{r}')]g(\bar{r},\bar{r}') + [\hat{n} \cdot \overline{E}(\bar{r}')]\nabla' g(\bar{r},\bar{r}') + [\hat{n} \times \overline{E}(\bar{r}')] \times \nabla' g(\bar{r},\bar{r}') \right\} \tag{21}$$

$$\overline{H}(\bar{r}) = \iint_{S'} dS' \left\{ -i\omega\epsilon[\hat{n} \times \overline{E}(\bar{r}')]g(\bar{r},\bar{r}') + [\hat{n} \cdot \overline{H}(\bar{r}')]\nabla' g(\bar{r},\bar{r}') + [\hat{n} \times \overline{H}(\bar{r}')] \times \nabla' g(\bar{r},\bar{r}') \right\} \tag{22}$$

Notice that in the Stratton-Chu formulas, normal components of $\overline{E}$ and $\overline{H}$ are required on the surface of S'.

The Franz formula can be derived from (21) and (22) to express $\overline{E}(\bar{r})$ and $\overline{H}(\bar{r})$ in terms of their tangential components on the surface S'. Taking curl of (22) yields, noting that $\nabla g(\bar{r},\bar{r}') = -\nabla' g(\bar{r},\bar{r}')$,

$$\begin{aligned} \nabla \times \overline{H}(\bar{r}) = &- i\omega\epsilon \nabla \times \iint_{S'} dS' [\hat{n} \times \overline{E}(\bar{r}')]g(\bar{r},\bar{r}') \\ &- \nabla \times \iint_{S'} dS' [\hat{n} \cdot \overline{H}(\bar{r}')]\nabla g(\bar{r},\bar{r}') \\ &+ \nabla \times \iint_{S'} dS' \nabla \times [\hat{n} \times \overline{H}(\bar{r}')]g(\bar{r},\bar{r}') \end{aligned} \tag{23}$$

The second term on the right-hand side of (23) vanishes and the curl operator inside the integral of the third term can be taken out of the integrand. Making use of Maxwell's equation $\nabla \times \overline{H} = -i\omega\epsilon\overline{E}$, we obtain

$$\begin{aligned} \overline{E}(\bar{r}) = &\nabla \times \iint_{S'} dS' [\hat{n} \times \overline{E}(\bar{r}')]g(\bar{r},\bar{r}') \\ &+ \frac{i}{\omega\epsilon} \nabla \times \nabla \times \iint_{S'} dS' [\hat{n} \times \overline{H}(\bar{r}')]g(\bar{r},\bar{r}') \end{aligned} \tag{24}$$

By duality

$$\begin{aligned} \overline{H}(\bar{r}) = &\nabla \times \iint_{S'} dS' [\hat{n} \times \overline{H}(\bar{r}')]g(\bar{r},\bar{r}') \\ &- \frac{i}{\omega\mu} \nabla \times \nabla \times \iint_{S'} dS' [\hat{n} \times \overline{E}(\bar{r}')]g(\bar{r},\bar{r}') \end{aligned} \tag{25}$$

The formulas are now dependent only on tangential components of the electric and magnetic fields on the surface S'.

b. Scalar Formulation

The Kirchhoff scalar formula for diffraction can be conveniently obtained from the Stratton-Chu formula (21). In view of the fact that $i\omega\mu\overline{H}(\bar{r}') = \nabla' \times \overline{E}(\bar{r}')$ and $(\hat{n} \times \overline{E}) \times \nabla' g = \overline{E}\hat{n} \cdot \nabla' g - \hat{n}\overline{E} \cdot \nabla' g$, (21) becomes

$$\begin{aligned}\overline{E}(\bar{r}) &= \oiint_{S'} dS' \left\{[\hat{n} \times (\nabla' \times \overline{E})]g + [\hat{n} \cdot \overline{E}]\nabla' g + \overline{E}\hat{n} \cdot \nabla' g - \hat{n}\overline{E} \cdot \nabla' g\right\} \\ &= \oiint_{S'} dS' \left\{\overline{E}\hat{n} \cdot \nabla' g - g(\hat{n} \cdot \nabla')\overline{E}\right\} \\ &\quad + \oiint_{S'} dS' \left\{\nabla'(\hat{n} \cdot g\overline{E}) - \hat{n}\nabla' \cdot (g\overline{E})\right\} \qquad (26)\end{aligned}$$

In arriving at the last term we use $\nabla' \cdot \overline{E}(\bar{r}') = 0$. The last surface integral in the above equation is zero owing to the fact that the ith component of the two surface integrals can be written as

$$\left[\oiint_{S'} dS' \nabla'(\hat{n} \cdot g\overline{E})\right]_i = \oiint_{S'} dS' \partial'_i n_j g E_j = \iiint_{V'} dV' \partial'_i \partial'_j g E_j \qquad (27)$$

and

$$\left[\oiint_{S'} dS' \hat{n}\nabla' \cdot (g\overline{E})\right]_i = \oiint_{S'} dS' n_i \partial_j g E_j = \iiint_{V'} dV' \partial'_i \partial'_j g E_j \qquad (28)$$

The last equality in (27) and (28) is the result of the generalized Gauss' theorem in tensor calculus. Thus

$$\overline{E}(\bar{r}) = \oiint_{S'} dS' \left\{\overline{E}(\bar{r}') \frac{\partial g(\bar{r}, \bar{r}')}{\partial n} - g(\bar{r}, \bar{r}') \frac{\partial \overline{E}(\bar{r}')}{\partial n}\right\} \qquad (29)$$

where $\partial/\partial n = \hat{n} \cdot \nabla'$. This is the vector form of the Kirchhoff formula for diffraction.

The scalar form of the Kirchhoff formula for diffraction is obtained by assuming that the electric field has only one cartesian component, say $\overline{E}(\bar{r}) = \hat{y}U(\bar{r})$. It follows that

$$U(\bar{r}) = \oiint_{S'} dS' \left\{U(\bar{r}') \frac{\partial g(\bar{r}, \bar{r}')}{\partial n} - g(\bar{r}, \bar{r}') \frac{\partial U(\bar{r}')}{\partial n}\right\} \qquad (30)$$

This is the most popular formula for diffraction used in physical optics. Clearly, when other components of the electric field become important, the formula breaks down. Thus in the case of diffraction of a linearly polarized plane wave by an aperture, for instance, the formula is valid only along the paraxial direction of observation i.e. near the axis perpendicular to the aperture. Away from the paraxial directions, the diffracted field components will no longer be in the same direction as the linearly polarized aperture field.

We have seen the limitations of the scalar theory from the point of view of the dyadic and the vector theories. However, the scalar diffraction theory is complete and consistent by itself. Equation (30) can be derived from the scalar equation

$$(\nabla^2 + k^2)U(r) = 0 \tag{31}$$

by using Green's theorem

$$U\nabla^2 g - g\nabla^2 U = \nabla \cdot [U\nabla g - g\nabla U] \tag{32}$$

Integrating (32) over the volume V enclosed by the surface S' and the surface at infinity and making use of (4) we obtain the Kirchhoff formula (30). Note that $\hat{n}$ is normal to surface S' and points into the region V and that the surface integral at infinity vanishes owing to the Sommerfeld radiation condition.

The normal derivative of the scalar Green's function in (30) can be written as

$$\frac{\partial g}{\partial n} = \hat{n} \cdot \nabla' g = -\left(ik - \frac{1}{R}\right)\frac{e^{ikR}}{4\pi R}(\hat{n} \cdot \hat{R})$$

where $\hat{R}$ represents the unit vector in the direction of $\overline{R} = \bar{r} - \bar{r}'$. Substituting in (30) we find

$$U(\bar{r}) = -\oint\!\!\!\!\!\!\int_{S'} dS' \left\{\left(ik - \frac{1}{R}\right)U(\bar{r}')(\hat{n} \cdot \hat{R}) + \frac{\partial U(\bar{r}')}{\partial n}\right\}\frac{e^{ikR}}{4\pi R} \tag{33}$$

The scalar wave amplitude $U(\bar{r})$ at $\bar{r}$ is thus expressed as a superposition of contributions from the elements of surface dS' with source strength expressed by the term within braces in (33).

We now consider the problem of diffraction by an aperture on an infinite plane surface [Fig. 5.3.2]. In order to apply (33), the so-called Kirchhoff approximations are made: (i) the values of $U(\bar{r}')$ and

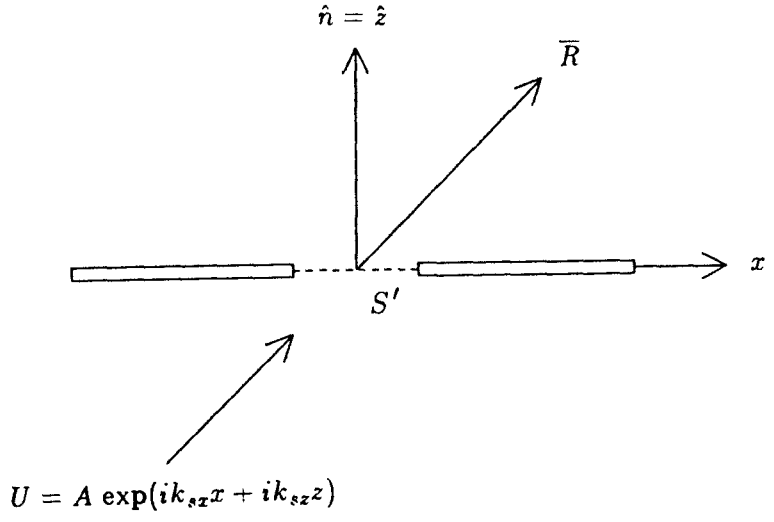


Figure 5.3.2 Diffraction by an aperture.

$\partial U(\bar{r}')/\partial n$ vanish everywhere on S' except at the aperture and (ii) the values of $U(\bar{r}')$ and $\partial U(\bar{r}')/\partial n$ at the aperture are equal to those of the incident wave in the absence of S'. Suppose that the aperture is illuminated by a plane wave

$$U(\bar{r}') = A \exp[ik_{sx}x' + ik_{sz}z'] \tag{34}$$

We have

$$\frac{\partial U}{\partial n} = \hat{n} \cdot \nabla' U(\bar{r}') = i(\hat{n} \cdot \bar{k}_s) U(\bar{r}') \tag{35}$$

where $\bar{k}_s = \hat{x}k_{sx} + \hat{z}k_{sz}$. Furthermore assume that the frequency is high such that the $1/R$ term in (33) is negligible in comparison with ik. In view of (35), we obtain from (33)

$$U(\bar{r}) = -ik \iint_{aperture} dS' \left\{ (\hat{n} \cdot \hat{R}) + (\hat{n} \cdot \hat{k}_s) \right\} U(\bar{r}') \frac{e^{ikR}}{4\pi R} \tag{36}$$

where $\hat{k}_s = \bar{k}_s/k$. The term $\{(\hat{n} \cdot \hat{R}) + (\hat{n} \cdot \hat{k}_s)\}/2$ is known as the *obliquity factor* which is always positive and less than unity. When the incident wave is normal to the aperture, the obliquity factor simply becomes $(1 + \cos\theta)/2$. If we view the integral as a superposition of

contributions from the secondary sources, then the obliquity factor effectively imposes an anisotropic directivity pattern on each element source point. Historically in the construction of the diffraction patterns with Huygens' wavelet concept, Fresnel found that the amplitude factor $ik(\hat{n}\cdot\hat{R}+\hat{n}\cdot\hat{k}_s)$ had to be postulated in order to get accurate results. It was Kirchhoff who showed that the amplitude factor is in fact a direct consequence of the scalar wave theory.

In the Kirchhoff approximation both $U(\bar{r}')$ and its normal derivative $\partial U(\bar{r}')/\partial n$ are prescribed over the surface S'. Rigorously the uniqueness theorem requires only $U(\bar{r}')$ or $\partial U(\bar{r}')/\partial n$ at every point on S' and not both. Sommerfeld circumvented the difficulty by choosing different scalar Green's functions $g(\bar{r}\cdot\bar{r}')$ such that on the plane surface S', where $g(\bar{r},\bar{r}')$ or $\partial g(\bar{r},\bar{r}')/\partial n$ vanishes and we have to specify only $\partial U(\bar{r}')/\partial n$ or $U(\bar{r}')$.

5.4 Fresnel and Fraunhofer Diffraction

As an example of the application of Huygens' principle, we consider the diffraction of a plane wave by a two-dimensional aperture. Let the plane wave be linearly polarized in the $\hat{y}$ direction and normally incident upon the aperture in the $x-y$ plane from the region $z<0$. In the half-space $z>0$, the wave is diffracted [Fig. 5.4.1]. To solve for the diffracted field, we use the results derived from the Huygens' principle,

$$\overline{E}(\bar{r})=\iint_{S'} dS'\left\{i\omega\mu\,\overline{\overline{G}}(\bar{r},\bar{r}')\cdot\overline{J}_s(\bar{r}')-\nabla\times\overline{\overline{G}}(\bar{r},\bar{r}')\cdot\overline{M}_s(\bar{r}')\right\} \quad (1a)$$

$$\overline{H}(\bar{r})=\iint_{S'} dS'\left\{i\omega\epsilon\,\overline{\overline{G}}(\bar{r},\bar{r}')\cdot\overline{M}_s(\bar{r}')+\nabla\times\overline{\overline{G}}(\bar{r},\bar{r}')\cdot\overline{J}_s(\bar{r}')\right\} \quad (1b)$$

We assume that at the aperture there is an electric current sheet $\overline{J}_s=2\hat{n}\times\overline{H}$ and no magnetic current sheet:

$$\overline{J}_s=2\hat{z}\times\overline{H}(\bar{r}')=-\hat{y}\frac{2}{\eta}E(x') \quad (2)$$

where $E(x')$ describes the aperture field distribution.

With the assumed surface current in (2),

$$\overline{E}(\bar{r})=-\frac{2}{\eta}\iint_{S'} dS'\,i\omega\mu\,\overline{\overline{G}}(\bar{r},\bar{r}')\cdot\hat{y}E(x')$$

The problem is two-dimensional, independent of the y coordinate, where $\partial/\partial y = 0$. We use the two-dimensional dyadic Green's function

$$\overline{\overline{G}}(\bar{r},\bar{r}') = \left(\overline{\overline{I}} + \frac{1}{k^2}\nabla\nabla\right) g(\bar{\rho},\bar{\rho}')$$

where

$$g(\bar{\rho},\bar{\rho}') = \frac{i}{4} H_0^{(1)}\left(k\sqrt{(x-x')^2+z^2}\right)$$

is the two-dimensional scalar Green's function. We find, recognizing that $\nabla\nabla\cdot\hat{y} = 0$ as $\partial/\partial y = 0$,

$$\overline{E}(x,z) = \hat{y}\frac{k}{2}\int_{-\infty}^{\infty} dx'\, E(x') H_0^{(1)}\left(k\sqrt{(x-x')^2+z^2}\right) \tag{3}$$

In the far-field zone, we use the asymptotic form for the Hankel function $H_0^{(1)}(\xi) \approx (2/i\pi\xi)^{1/2} e^{i\xi}$. We also assume $z \gg |x-x'|$ and expand

$$k\sqrt{(x-x')^2+z^2} \approx kz\left[1 + \frac{1}{2}\left(\frac{x-x'}{z}\right)^2 + \cdots\right] \tag{4}$$

Equation (3) becomes

$$\overline{E}(x,z) = \hat{y}\frac{k}{2}\sqrt{\frac{2}{i\pi kz}}\, e^{ikz}\int_{-\infty}^{\infty} dx'\, E(x') e^{ik(x-x')^2/2z} \tag{5}$$

where we use the first two terms of expansion (4) in the exponent and only the first term in the denominator of the Hankel function expansion.

The expansion (4) which gives rise to (5) is referred to as the Fresnel approximation. Equation (5) is conveniently evaluated in terms of tabulated functions known as Fresnel integrals defined by

$$F(w) = \int_0^w dt\, e^{i\pi t^2/2} = C(w) + iS(w) \tag{6}$$

We notice the symmetry relation

$$C(w) = \int_0^w dt \cos\left(\frac{\pi t^2}{2}\right) = -C(-w) \tag{7}$$

$$S(w) = \int_0^w dt \sin\left(\frac{\pi t^2}{2}\right) = -S(-w) \tag{8}$$

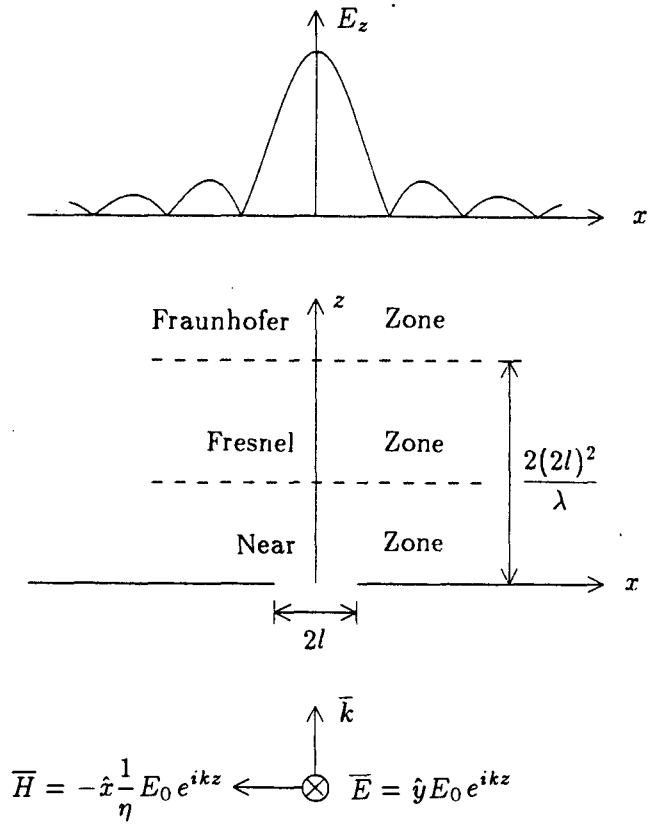


Figure 5.4.1 Diffraction field in the Fraunhofer zone.

For small values of w, the integral can be expanded in power series of w to carry out the integration. As $w \to 0$, we have

$$F(w) \approx w$$

For large values of w, the integral can be expanded in series of inverse powers of w. As $w \to \infty$, we have

$$F(\infty) = \int_0^\infty dt\, e^{i\pi t^2/2} = \sqrt{2/\pi}e^{i\pi/4}\int_0^\infty dx\, e^{-x^2} = \frac{1}{2} + i\frac{1}{2}$$

The values are shown in Figure 5.4.2a. When they are plotted on the two-dimensional space formed by $C(w)$ and $S(w)$, the trajectory of $C(w) + iS(w)$ forms the Cornu spiral as shown in Figure 5.4.2b.

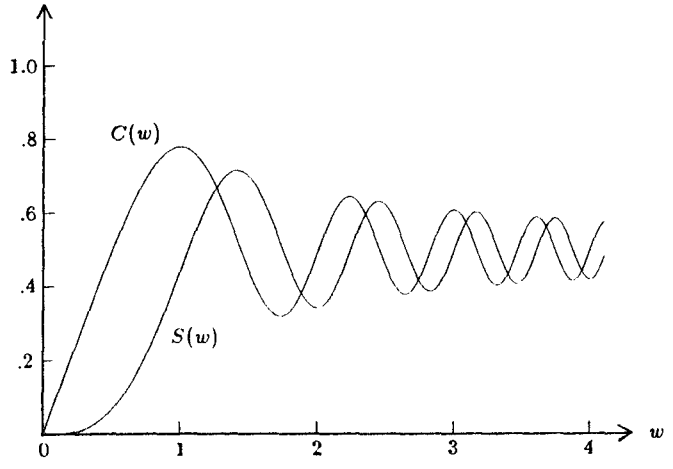


Figure 5.4.2a Fresnel integrals $C(w)$ and $S(w)$.

Note that

$$\frac{dC(w)}{dw} = \cos\left(\frac{\pi}{2}w^2\right)$$
$$\frac{dS(w)}{dw} = \sin\left(\frac{\pi}{2}w^2\right)$$

The slope of the Cornu spiral is seen to be

$$\tan\theta = \frac{dS/dw}{dC/dw} = \tan\left(\frac{\pi}{2}w^2\right) \tag{9}$$

The angle θ increases monotonically with w^2. The tangent to the curve is vertical when w^2 is an odd integer and horizontal when w^2 is an even integer.

The incremental length along the curve is

$$(dC)^2 + (dS)^2 = \left[\left(\frac{dC}{dw}\right)^2 + \left(\frac{dS}{dw}\right)^2\right] dw^2 = dw^2 \tag{10}$$

Thus w is the length of the curve from the origin.

To illustrate the use of the Cornu spiral, we first consider the diffraction by a half-space aperture. With the integration limit from 0 to ∞, the electric field in (6) takes the form

$$\overline{E}(x,z) = \hat{y}\frac{E_0}{\sqrt{2i}}e^{ikz}D(x)$$

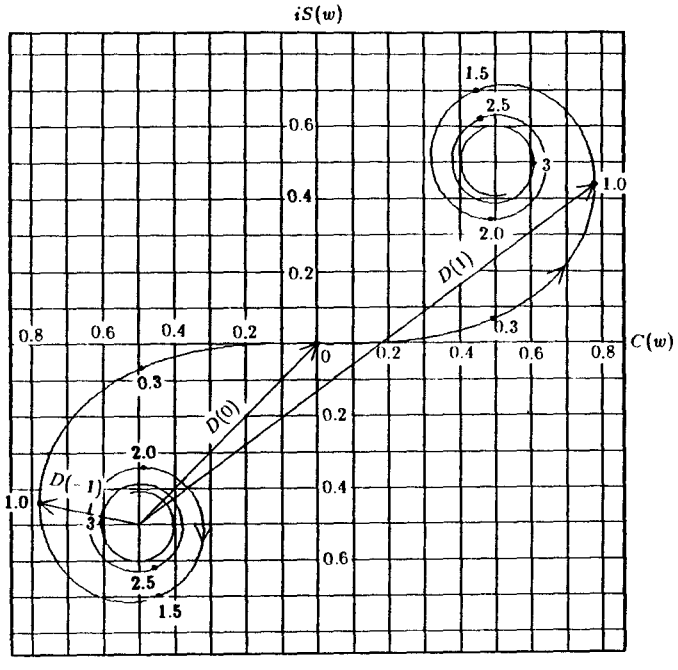


Figure 5.4.2b Cornu spiral.

where

$$D(x) = \sqrt{\frac{k}{\pi z}} \int_0^\infty dx'\, e^{ik(x'-x)^2/2z} = \int_0^\infty dt\, e^{i\pi t^2/2} - \int_0^{-x} dt\, e^{i\pi t^2/2}$$

$$= F(\infty) - F(-\sqrt{k/\pi z}\, x)$$

$$= \frac{1}{2} - C(-\sqrt{k/\pi z}\, x) + i\left[\frac{1}{2} - S(-\sqrt{k/\pi z}\, x)\right]$$

The magnitude of $\overline{E}$ is proportional to the magnitude of $D(x)$. We note the following special values for $D(x)$:

$$D(-\infty) = 0$$

$$D(0) = \frac{1}{2} + i\frac{1}{2}$$

$$D(\infty) = 1 + i$$

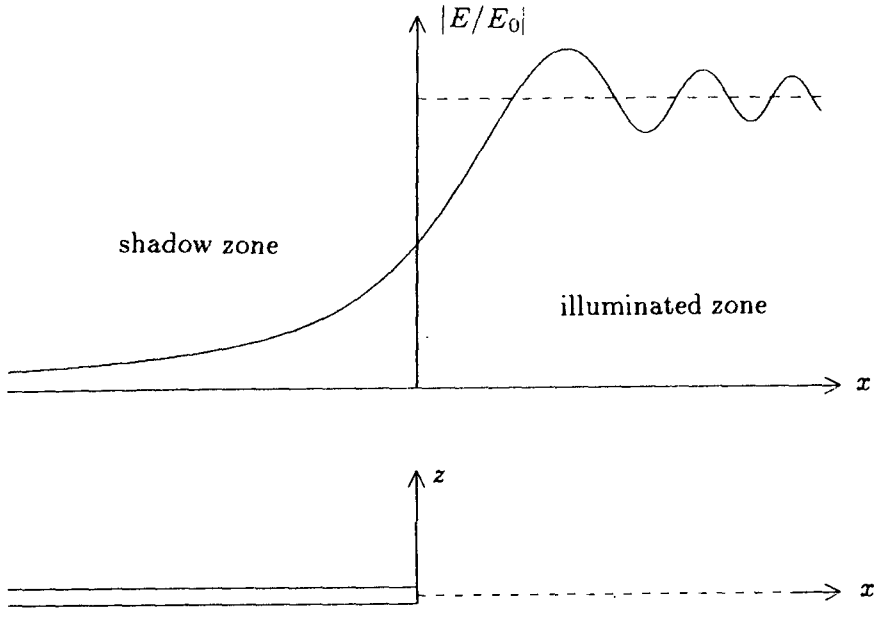


Figure 5.4.3 Diffraction by a half-plane aperture.

On the Cornu spiral [Fig. 5.4.2a], we may use the center of the lower left spiral as the reference point for $D(-\infty)$. Then the distance from $D(-\infty)$ to $x = 0$ is the magnitude for $D(0)$. The corresponding point for $D(\infty)$ is at the center of the upper right spiral. The distance from $D(-\infty)$ is the magnitude for $D(\infty)$. When we move from deep in the shadow zone for $x < 0$ towards the illuminated zone for $x > 0$, the magnitude of $D(x)$ is the distance from the reference point $D(-\infty)$ to a point corresponding to the value of x on the spiral. Thus $|D(x)|$ first increases its value monotonically until $x = 0$ and becomes oscillatory and approaches the final value of $\sqrt{2}$. If we normalize the value by $\sqrt{2}$ then the final value is 1 and the value at $x = 0$ is $1/2$ [Fig. 5.4.3].

We now consider the aperture with width $2l$ in the $\hat{x}$ direction [Fig. 5.4.1] and infinite in the $\hat{y}$ direction. The aperture surface current is, according to (2),

$$\overline{J}_s = -\hat{y}\frac{2E_0}{\eta}U(l - |x'|)$$

where E_0 is the amplitude of the incident wave, $U(l-|x'|)$ is a unit step function that is unity for $|x'| \leq l$ and zero for $|x'| \geq l$. This assumption is justifiable, for instance, in the case of optical diffraction by an opaque screen. In the case of perfectly conducting screens, however, the boundary condition of zero tangential electric field at $x = \pm l$ is not enforced.

At a distance z in the $x-z$ plane, we have

$$\overline{E}(x,z) = \hat{y}\frac{E_0}{\sqrt{2i}}e^{ikz}D(x)$$

with

$$\begin{aligned} D(x) &= F\left[(l-x)\sqrt{k/\pi z}\right] - F\left[-(l+x)\sqrt{k/\pi z}\right] \\ &= F\left[(l-x)\sqrt{k/\pi z}\right] + F\left[(l+x)\sqrt{k/\pi z}\right] \end{aligned}$$

The second equality follows from the symmetry relation for $F(w)$ in (8). Clearly $|D(x)|$ is symmetrical with respect to $x = 0$. As seen from (10), it corresponds to the distance between the two end points of an arc of constant length proportional to $2l$. At $x = 0$, the magnitude $|D(0)|$ is the greatest. As x increases, $|D(x)|$ first decreases and then becomes oscillatory. For very large z, we move into the Fraunhofer zone where further simplifications can be made.

When the observation point is very far from the finite aperture, the Fresnel approximation for the two-dimensional slit diffraction problem can be further simplified. We have

$$\begin{aligned} k\sqrt{(x-x')^2+z^2} &\approx kz\left[1+\frac{1}{2}\left(\frac{x-x'}{z}\right)^2\right] \\ &\approx kz + \frac{kx^2}{2z} - \frac{kxx'}{z} \end{aligned} \tag{11}$$

This is known as the *Fraunhofer approximation*. The electric field in the Fraunhofer zone due to a plane wave normally incident on a slit with width $2l$ becomes

$$\overline{E}(x,z) = \hat{y}\frac{k}{2}\sqrt{\frac{2}{i\pi kz}}e^{ikz+ikx^2/2z}\int_{-\infty}^{\infty} dx'\, E_0\, U(l-|x'|)e^{-i(kx/z)x'} \tag{12}$$

The integral is a Fourier transform of the step function, and

$$\overline{E}(x,z) = \hat{y}klE_0\sqrt{\frac{2}{i\pi kz}}\frac{\sin(kxl/z)}{kxl/z}e^{ikz+ikx^2/2z}$$

Note that in (12) the source function $E_0U[l - |x'|]$ can be replaced by any general aperture field distribution. A general conclusion is that in the Fraunhofer zone, characterized by approximation (11), the observed field at a constant z is proportional to the Fourier transform of the aperture field as seen from (12). The quadrature phase term $e^{ikx^2/2z}$ in (12) accounts for the curvature of the curved phase front. In the Fresnel approximation the inclusion of the quadrature phase correction term $e^{ikx'^2/2z}$ at the aperture leads to the use of Fresnel integrals. In the radiation field approximation, which is not restricted to paraxial directions, the quadrature phase term in Fraunhofer formula is further neglected.

According to the distance from the aperture, the diffracted field can be divided into the near zone, the Fresnel zone, and the Fraunhofer zone [Fig. 5.4.1]. The separation between the Fresnel zone and the Fraunhofer zone may be taken at $z_F = 2(2l)^2/\lambda$. For an aperture of dimension 1 cm and wavelength 0.63 μm, the Fraunhofer zone begins at $z_F \approx 320\,\text{m}$, a rather stringent constraint for Fourier optics experiments. The use of a convergent lens may eliminate this problem. Under paraxial approximation, the effect of a lens amounts to giving a wave transmitting through the lens a phase factor proportional to $e^{-ikx'^2/2f}$, where f is the focal length. If a lens is placed immediately in front of the aperture, we find that at $z = f$ the Fresnel diffraction formula reduces to the Fraunhofer formula under the Fresnel approximation. Thus when a screen is placed a focal length away from the lens, the Fourier transform of the aperture field is observed.

In general, when the observation point is very far from the aperture, the diffracted field is equivalent to the radiation field due to an aperture antenna with a surface current distribution. The radiation field approximation takes the form

$$|\bar{r} - \bar{r}'| \approx r - \hat{r} \cdot \bar{r}'$$

where r is the magnitude of $\bar{r}$ and $\hat{r}$ is a unit vector along $\bar{r}$, with $\bar{r} = \hat{r}r$. The radiated field is also in the direction of $\bar{r}$, i.e., $\bar{k} = \hat{r}k$. For three-dimensional problems in free space, the dyadic Green's function is approximated by

$$\bar{\bar{G}}(\bar{r}, \bar{r}') = \left[\bar{\bar{I}} + \frac{1}{k^2}\nabla\nabla\right] \frac{e^{ik|\bar{r}-\bar{r}'|}}{4\pi\,|\bar{r} - \bar{r}'|} \approx \frac{e^{ikr}}{4\pi r}\left[\bar{\bar{I}} - \hat{r}\hat{r}\right] e^{-i\bar{k}\cdot\bar{r}'}$$

where the del operator ∇ is replaced by $i\bar{k}$. Huygens' principle can be written in the following form:

$$\bar{E}(\bar{r}) \approx \frac{i\omega\mu e^{ikr}}{4\pi r} \oiint_{S'} dS' e^{-i\bar{k}\cdot\bar{r}'} \cdot \left\{ \left[\bar{\bar{I}} - \hat{r}\hat{r}\right] \cdot \bar{J}_s(\bar{r}') - \frac{1}{\eta}\hat{r} \times \bar{M}_s(\bar{r}') \right\} \tag{14a}$$

$$\bar{H}(\bar{r}) \approx \frac{ik\, e^{ikr}}{4\pi r} \oiint_{S'} dS' e^{-i\bar{k}\cdot\bar{r}'} \cdot \left\{ \hat{r} \times \bar{J}_s(\bar{r}') + \frac{1}{\eta}\left[\bar{\bar{I}} - \hat{r}\hat{r}\right] \cdot \bar{M}_s(\bar{r}') \right\} \tag{14b}$$

The two equations are seen to be duals of each other with the replacements $\bar{E} = \eta\bar{H}$, $\bar{H} = -\bar{E}/\eta$, $\bar{J}_s = \bar{M}_s/\eta$ and $\bar{M}_s = -\eta\bar{J}_s$. Observe that the first term in (14a) is identical to the radiation field calculated previously except that instead of volume current densities we are integrating surface current densities over the surface enclosing the source.

As the first example, consider the radiation due to a rectangular aperture antenna with a constant surface current sheet

$$\bar{J}_s = \hat{x}J_s$$

Consider the approximation

$$\begin{aligned} |\bar{r} - \bar{r}'| &= (z^2 + x^2 + y^2 - 2xx' - 2yy' + x'^2 + y'^2)^{1/2} \\ &= (x^2 + y^2 + z^2)^{1/2} - \frac{xx'}{r} - \frac{yy'}{r} + \delta \end{aligned}$$

The error term may be approximated by the quadrature phase

$$\delta \approx \frac{(D_x/2)^2}{2r} + \frac{(D_y/2)^2}{2r}$$

where D_x and D_y are the dimensions of the aperture. Let us apply an arbitrary criterion restricting the phase error due to each dimension to be less than $\pi/8$. Letting D denote either D_x or D_y,

$$\frac{\pi}{8} > \frac{k(D/2)^2}{2r} = \frac{2\pi D^2}{8\lambda r}$$

Thus

$$r > \frac{2D^2}{\lambda}$$

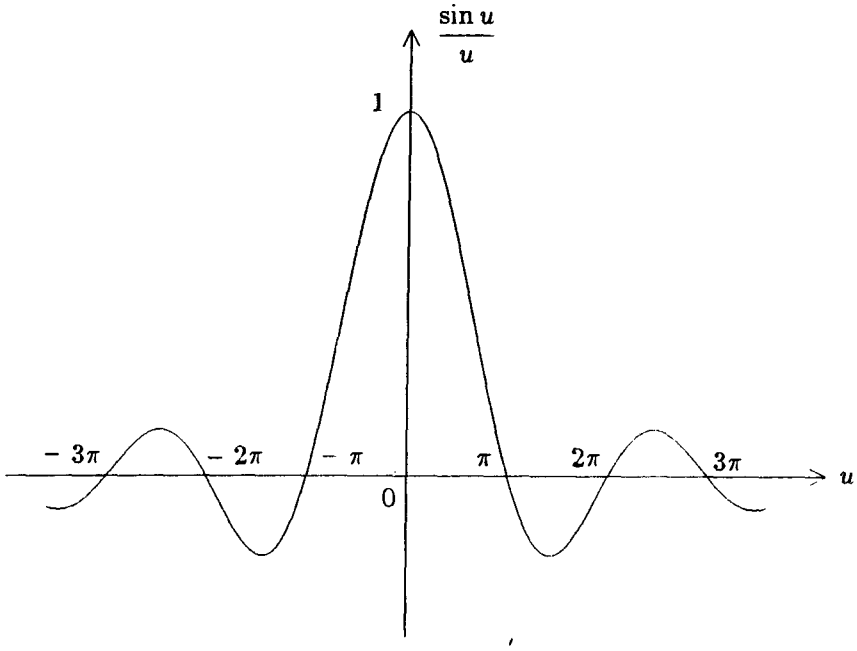


Figure 5.4.4 Plot of $\sin u/u$.

Beyond this distance, we are in the Fraunhofer zone. Making use of the formula in the radiation field, the electric field vector is found from (3a) to be

$$E_x = \frac{i\omega\mu J_s e^{ikr}}{4\pi r} \int_{-D_x/2}^{D_x/2} dx' \, e^{-ikx' \sin\theta\cos\phi} \int_{-D_y/2}^{D_y/2} dy' \, e^{-iky' \sin\theta\sin\phi}$$

$$= \frac{i\omega\mu J_s e^{ikr}}{4\pi r} D_x D_y \frac{\sin\left(\dfrac{kD_x}{2}\sin\theta\cos\phi\right)}{\dfrac{kD_x}{2}\sin\theta\cos\phi} \frac{\sin\left(\dfrac{kD_y}{2}\sin\theta\sin\phi\right)}{\dfrac{kD_y}{2}\sin\theta\sin\phi}$$

A plot of $(\sin u)/u$ is shown in Figure 5.4.4. The value at $u = \pm 0.44\pi$ is approximately $1/\sqrt{2}$. The half-power beamwidth (HPBW) is the angle at which the radiated power is half of its value at $u = 0$. The maximum radiation intensity is

$$r^2 Re\{S\} = \frac{1}{2\eta}\left[\omega\mu \frac{J_s}{4\pi} D_x D_y\right]^2$$

The total radiated power can be calculated by integrating the field over the aperture. As $\overline{E} = -\hat{x}\eta J_s/2$ and $\overline{H} = -\hat{y}J_s/2$, we find

$$P_t = \frac{\eta}{8} J_s^2 D_x D_y$$

and the directivity

$$D = \frac{4\pi r^2 Re\{S\}}{P_t} = \frac{k^2}{\pi} D_x D_y$$

It is important to note that the radiation characteristics are valid in the far-field zone commonly defined by $r > 2D^2/\lambda$.

As another example, consider the diffraction of a plane wave normally incident on a circular aperture. We shall not restrict observation points to be near the z axis. At the aperture we use Equivalent Problem 3 for a plane wave as discussed in 5.1. We assume a magnetic current sheet with $\overline{M}_s = -2\hat{n} \times \overline{E}$ and no electric current sheet. We find from Huygens' principle

$$\begin{aligned} \overline{E}(\overline{r}) &= 2 \oiint_{S'} dS' \nabla \times \overline{\overline{G}}(\overline{r}, \overline{r}') \cdot [\hat{n} \times \overline{E}(\overline{r}')] \\ &= 2\nabla \times \oiint_{S'} dS' \hat{n} \times \overline{E}(\overline{r}') \frac{e^{ik|\overline{r}-\overline{r}'|}}{4\pi |\overline{r} - \overline{r}'|} \end{aligned} \tag{15}$$

where S' denotes the area of the circular aperture.

In the far-field zone, the observation point is so remote from the aperture that all wave vectors originating from the aperture are essentially parallel and we have $k|\overline{r} - \overline{r}'| \approx kr - \overline{k} \cdot \overline{r}'$. The integral in (15) becomes

$$\overline{E}(\overline{r}) \approx \frac{ie^{ikr}}{2\pi r} \overline{k} \times \oiint_{S'} dS' \hat{n} \times \overline{E}(\overline{r}') e^{-i\overline{k} \cdot \overline{r}'}$$

Let the incident plane wave be polarized in the $\hat{x}$ direction, $\overline{k} = \hat{x}k \sin\theta \cos\phi + \hat{y}k \sin\theta \sin\phi + \hat{z}k \cos\theta$. We have

$$\begin{aligned} \overline{E}(\overline{r}) &= \frac{iE_0 e^{ikr}}{2\pi r} (\overline{k} \times \hat{y}) \int_0^R d\rho' \int_0^{2\pi} \rho' d\phi' \, e^{-ik\rho' \sin\theta \cos(\phi - \phi')} \\ &= (\hat{z} \sin\theta \cos\phi - \hat{x} \cos\theta) \frac{ikE_0 e^{ikr}}{r} \int_0^R \rho' d\rho' J_0(k\rho' \sin\theta) \\ &= (\hat{z} \sin\theta \cos\phi - \hat{x} \cos\theta) \frac{ikR^2 E_0 e^{ikr}}{r} \frac{J_1(kR \sin\theta)}{kR \sin\theta} \end{aligned}$$

where E_0 is the amplitude of the incident wave and R is the radius of the aperture.

In the previous examples, we assume the aperture to be a hole on an opaque screen and the equivalent source is equal to the part of the surface current sheet due to the original plane wave. We now consider a rectangular waveguide carrying a TE_{10} mode opening into free space. By the equivalence principle, we may approximate the aperture field at the waveguide opening by a current sheet with

$$\overline{J}_s(\vec{r}') = -\hat{y}\frac{2E_0}{\eta}\cos\frac{\pi x'}{a}$$

for $-a/2 < x' < a/2$ and $-b/2 < y' < b/2$. The far field is the Fourier transform of the aperture source

$$\overline{E} = -\hat{\theta}i\omega\mu\frac{2E_0}{\eta}\cos\theta\frac{e^{ikr}}{4\pi r}\int_{-b/2}^{b/2}dy'\int_{-a/2}^{a/2}dx'\cos\frac{\pi x'}{a}e^{-ik_x x'-ik_y y'}$$

$$= -\hat{\theta}i\omega\mu\frac{4E_0}{\eta}\cos\theta\frac{e^{ikr}}{4\pi r}\frac{\sin k_y\frac{b}{2}}{k_y}$$

$$\left\{\frac{\sin\left[\left(k_x+\frac{\pi}{a}\right)\frac{a}{2}\right]}{k_x+\frac{\pi}{a}}+\frac{\sin\left[\left(k_x-\frac{\pi}{a}\right)\frac{a}{2}\right]}{k_x-\frac{\pi}{a}}\right\}$$

where $k_x = k\sin\theta\cos\phi$ and $k_y = k\sin\theta\sin\phi$.

5.5 Reaction and Reciprocity

a. Reaction

Consider a time-harmonic source a, denoted by $\overline{J}_a$ and $\overline{M}_a$, in a field $\overline{E}_b$ and $\overline{H}_b$ produced by source b, denoted by $\overline{J}_b$ and $\overline{M}_b$. The interaction of source a with field b can be characterized by $< a, b >$, defined as

$$< a, b > \equiv \iiint_V dV\ (\overline{J}_a\cdot\overline{E}_b - \overline{M}_a\cdot\overline{H}_b) \tag{1}$$

Note that in the representation $< a, b >$ the first entry, a, is associated with the source and the second entry, b, with the field; whenever the source is zero, the reaction is zero. The integration extends over

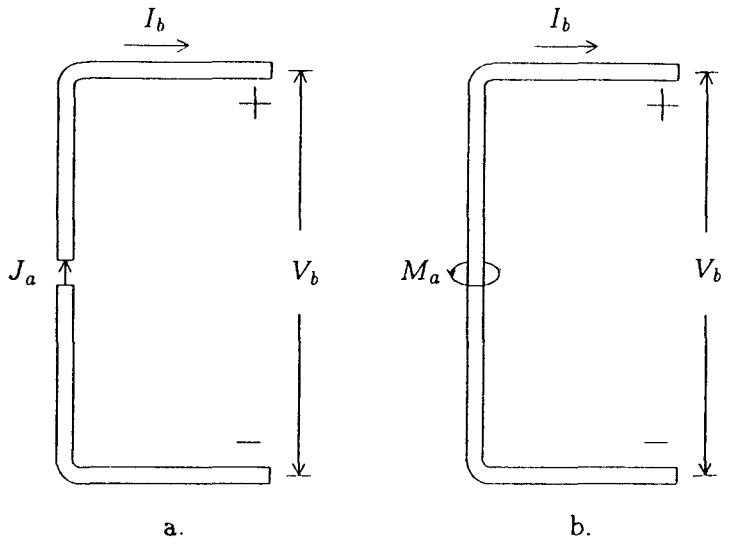


Figure 5.5.1 a. Reaction of a current source. b. Reaction of a voltage source.

the region containing source a, which is composed of volume current densities, as well as surface current densities. In the case of surface current densities, the volume integrals become surface integrals.

The reaction is a complex number and has the dimension of power. It is different from complex power in two respects; first, in the definition of power the current density is complex-conjugated; second, the reaction is defined for a source with respect to the field produced by another source. When the source is reacting to the field produced by itself, we have the self-reaction $< a, a >$.

To understand what reaction means physically, consider the case in which source a is a dipole, $\overline{J}_a = \overline{I}l\,\delta(\overline{r} - \overline{r}_0)$. Then the reaction

$$< a, b > = \overline{I}l \cdot \overline{E}_b(\overline{r}_0) \tag{2}$$

is proportional to the electric field in the direction of $\overline{I}$ as measured by the dipole. It is equal to the field strength produced by source b at the dipole when the dipole moment Il is unity. For instance, source b can be an antenna located at the origin.

As another example, consider the reactions of a current source and a voltage source in circuits [Fig. 5.5.1]. Let V_b and I_b be caused by

an unspecified source b. For the current source [Fig. 5.5.1a], we have

$$\begin{aligned} < a,b > &= \iiint_V dV\, \overline{J}_a \cdot \overline{E}_b \\ &= I_a \int d\bar{l} \cdot \overline{E}_b = -I_a V_b \end{aligned} \tag{3a}$$

For the voltage source [Fig. 5.5.1b], we have

$$\begin{aligned} < a,b > &= -\iiint_V dV\, \overline{M}_a \cdot \overline{H}_b \\ &= -V_a \int d\bar{l} \cdot \overline{H}_b = -V_a I_b \end{aligned} \tag{3b}$$

Thus, if we use a unit current source, the reaction $< a,b >$ is equal to the voltage V_b at source a due to source b. If we use a unit voltage source, the reaction $< a,b >$ is equal to the current I_b at source a due to source b.

b. Reciprocity

We define a system as reciprocal if, with respect to two sets of sources a and b,

$$< a,b > = < b,a > \tag{4}$$

Let the system be an isotropic medium; we shall show that isotropic media are reciprocal. We write Maxwell's equations with source a as

$$\nabla \times \overline{H}_a = -i\omega\epsilon \overline{E}_a + \overline{J}_a \tag{5a}$$

$$-\nabla \times \overline{E}_a = -i\omega\mu \overline{H}_a + \overline{M}_a \tag{5b}$$

and Maxwell's equations with source b as

$$\nabla \times \overline{H}_b = -i\omega\epsilon \overline{E}_b + \overline{J}_b \tag{6a}$$

$$-\nabla \times \overline{E}_b = -i\omega\mu \overline{H}_b + \overline{M}_b \tag{6b}$$

Forming the sum of (5a) dot-multiplied by $\overline{E}_b$ and (5b) dot-multiplied by $\overline{E}_a$, we arrive at

$$-\nabla \cdot (\overline{E}_b \times \overline{H}_a) = -i\omega\epsilon \overline{E}_a \cdot \overline{E}_b + \overline{J}_a \cdot \overline{E}_b - i\omega\mu \overline{H}_a \cdot \overline{H}_b + \overline{M}_b \cdot \overline{H}_a \tag{7a}$$

$$-\nabla \cdot (\overline{E}_a \times \overline{H}_b) = -i\omega\epsilon \overline{E}_a \cdot \overline{E}_b + \overline{J}_b \cdot \overline{E}_a - i\omega\mu \overline{H}_a \cdot \overline{H}_b + \overline{M}_a \cdot \overline{H}_b \tag{7b}$$

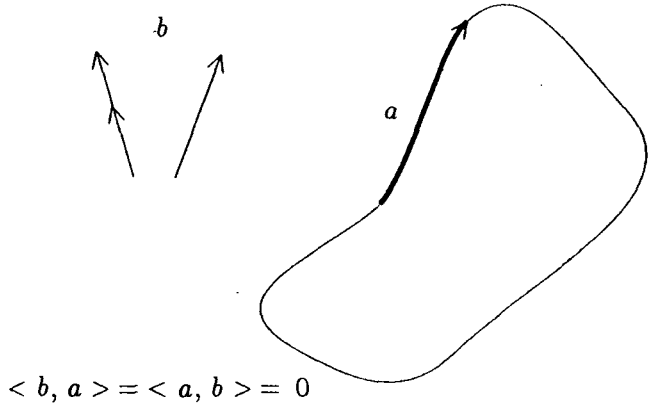


Figure 5.5.2 Zero field produced by current sheet impressed on perfect conductor.

Subtracting (7*b*) from (7*a*) and integrating, we obtain

$$< a,b > - < b,a > = \oiint_S d\overline{S} \cdot (\overline{E}_a \times \overline{H}_b - \overline{E}_b \times \overline{H}_a) \tag{8}$$

By definition, isotropic media are reciprocal provided that

$$\oiint_S d\overline{S} \cdot (\overline{E}_a \times \overline{H}_b - \overline{E}_b \times \overline{H}_a) = 0 \tag{9}$$

This statement is referred to as the *Lorentz reciprocity theorem.* When all sources and matter are of finite extent, we can extend the surface of the volume to infinity and the left-hand side of (8) contains all sources contributing to the reactions. At an infinite distance away from the source, the $\overline{E}$ and $\overline{H}$ fields are related by $\overline{H} = \hat{r} \times \overline{E}/\eta$ and $\hat{r} \cdot \overline{E} = \hat{r} \cdot \overline{H} = 0$. As a consequence, $\overline{E}_a \times \overline{H}_b - \overline{E}_b \times \overline{H}_a = 0$ and we find that the surface integral in (9) vanishes.

The reciprocity theorem that we have just proved is useful in many situations. We shall apply it in subsequent sections to derive stationary formulas in variational problems based on reciprocity. We now use it to prove a few simple assertions. First, consider an electric current sheet impressed on the surface of a perfect conductor. If the surface of the conductor is a plane, the image theorem assures us that no field is produced by the current. But what if the surface is not a plane? Does it seem reasonable to extend the result and claim that all electric

current sheets impressed on the surface of a conductor of any shape do not produce a field? The reciprocity theorem assures us that this is true by a simple argument. Let the impressed source by denoted as source a [Fig. 5.5.2]. Let there be a source b that measures the field produced by source a; for instance, source b can be a dipole antenna. Source b produces no tangential electric field along the surface of the conductor, because of the boundary conditions. The reaction of a and b is $< a,b >= 0$. By the reciprocity theorem,

$$< b,a >=< a,b >= 0$$

But $< b,a >$ is the field arising from the impressed source a as measured by source b, and b can be any arbitrary source. Therefore, source b measures no field, and this proves that impressed electric current sheets on the surface of a perfect conductor produce no field.

For two antennas operating as a transmitter and a receiver, we can regard the terminals of the two antennas as the terminals of a two-port network in circuit theory. We write

$$V_a = Z_{aa}I_a + Z_{ab}I_b \tag{10a}$$
$$V_b = Z_{ba}I_a + Z_{bb}I_b \tag{10b}$$

Suppose both antennas a and b are excited with terminal currents I_a and I_b. Since no magnetic sources are present, $< a,b >=< b,a >$ gives

$$\iiint_{V_a} dV\, \bar{J}_a \cdot \bar{E}_b = \iiint_{V_b} dV\, \bar{J}_b \cdot \bar{E}_a$$

For perfectly conducting antennas, the electric fields are zero over the antennas and we have

$$V_a^{oc} I_a = V_b^{oc} I_b \tag{11}$$

where

$$V_a^{oc} = -\int d\bar{l} \cdot \bar{E}_b$$

is the open circuit voltage at the terminal of antenna a due to the field produced by antenna b, and similarly V_b^{oc} is the open circuit voltage at the terminal of antenna b due to the field produced by antenna a. From (10) we have

$$V_a^{oc} = Z_{ab}I_b$$
$$V_b^{oc} = Z_{ba}I_a$$

It follows from (11) that

$$Z_{ab} = Z_{ba} \tag{12}$$

This is a direct consequence of the reciprocity principle. If a current source I excites antenna a, the open circuit voltage at the terminals of antenna b will be $V_b^{oc} = Z_{ba}I$ which is equal to $V_a^{oc} = Z_{ab}I$ at the terminals of antenna a when the same current source is used to excite antenna b.

We now prove that the receiving pattern of an antenna is identical to its radiation pattern. Let antenna a be the antenna under consideration, and antenna b be a test antenna, constructed so that it is omnidirectional. For both receiving and radiation patterns we are concerned with far fields. By the reciprocity theorem, we see that, in a direction where a as a transmitter radiates a weaker plane wave to b, a as a receiver also receives a weaker plane wave from b. Thus the radiation pattern and the receiving pattern of antenna a are identical. The gain $G_a(\theta,\phi)$ characterizes the radiation pattern of antenna a when it is acting as a transmitter. When a is acting as a receiver, we can define an effective area $A_a(\theta,\phi)$ to characterize its receiving pattern. From the proof we know that, for any antenna a, the effective area $A(\theta,\phi)$ is related to its gain $G(\theta,\phi)$ by a constant independent of the structure of the antenna. To determine this constant, we consider a dipole antenna with the gain function

$$G(\theta,\phi) = \frac{3}{2}\sin^2\theta \tag{13}$$

Let there be a plane wave incident upon the dipole from a direction making angle θ with the dipole axis along $\bar{l}$ [Fig. 5.5.3]. The power received by the antenna with a properly matched load Z_L is

$$P = \frac{(\bar{E}\cdot\bar{l})^2}{2(Z_L + Z_L^*)^2}R_r = \frac{E^2l^2\sin^2\theta}{8R_r} = \left[\frac{\pi}{k^2}\frac{3}{2}\sin^2\theta\right]S_{inc}$$

The last equality follows on using the dipole radiation resistance $R_r = (\eta/6\pi)k^2l^2$, and the incident power density $S_{inc} = E^2/2\eta$. Thus we define

$$A(\theta,\phi) = \frac{\pi}{k^2}G(\theta,\phi) \tag{14}$$

Thus the power received by the antenna is equal to the power per unit area of the incident wave times the effective area $A(\theta,\phi)$. Notice that

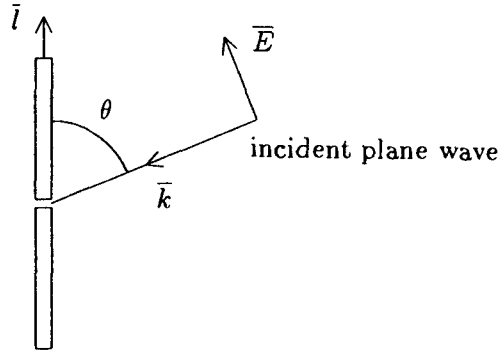


Figure 5.5.3 Dipole as receiver.

the proportionality factor, π/k^2, even though determined for a dipole antenna, is a universal constant independent of antenna structures.

c. Modified Reciprocity Theorem

In the preceding development we proved that isotropic media are reciprocal, and we applied the reciprocity theorem in various situations. We shall now examine the validity of the reciprocity theorem for a bianisotropic medium. Using the same procedure as illustrated in (5)–(9), we find

$$< a,b > - < b,a > = i\omega \iiint_V dV \, (\overline{E}_b \cdot \overline{D}_a - \overline{E}_a \cdot \overline{D}_b + \overline{H}_a \cdot \overline{B}_b - \overline{H}_b \cdot \overline{B}_a) \quad (15)$$

If the right-hand side is zero, the medium is reciprocal. The constitutive relations for bianisotropic media are

$$\overline{D} = \bar{\bar{\epsilon}} \cdot \overline{E} + \bar{\bar{\xi}} \cdot \overline{H} \quad (16a)$$

$$\overline{B} = \bar{\bar{\mu}} \cdot \overline{H} + \bar{\bar{\zeta}} \cdot \overline{E} \quad (16b)$$

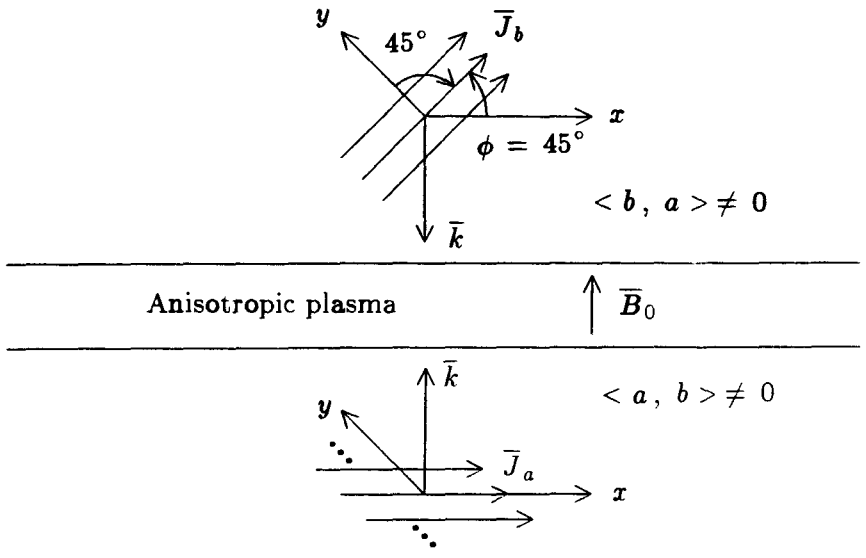


Figure 5.5.4 Transmission through an anisotropic plasma slab.

Inserting (16) into the right-hand side of (15), we obtain

$$< a,b > - < b,a >$$
$$= i\omega \iiint_V dV \left[\overline{E}_b \cdot (\overline{\overline{\epsilon}} - \overline{\overline{\epsilon}}^T) \cdot \overline{E}_a + \overline{H}_a \cdot (\overline{\overline{\mu}} - \overline{\overline{\mu}}^T) \cdot \overline{H}_b \right.$$
$$\left. + \overline{E}_b \cdot (\overline{\overline{\xi}} + \overline{\overline{\zeta}}^T) \cdot \overline{H}_a + \overline{H}_a \cdot (\overline{\overline{\zeta}} + \overline{\overline{\xi}}^T) \cdot \overline{E}_b \right]$$

Thus the medium will be reciprocal if

$$\overline{\overline{\epsilon}} = \overline{\overline{\epsilon}}^T \tag{17a}$$
$$\overline{\overline{\mu}} = \overline{\overline{\mu}}^T \tag{17b}$$
$$\overline{\overline{\xi}} = -\overline{\overline{\zeta}}^T \tag{17c}$$

These are the conditions for a medium to be reciprocal. Consequently, isotropic media are reciprocal, and anisotropic media with symmetrical permittivity and permeability tensors are reciprocal. Bianisotropic media that satisfy (17*a*), (17*b*), and the symmetry condition $\overline{\overline{\xi}} = \overline{\overline{\zeta}}^+$ are reciprocal if $\overline{\overline{\xi}}$ and $\overline{\overline{\zeta}}$ are purely imaginary matrices.

An anisotropic plasma is an example of a nonreciprocal medium. It possesses a permittivity tensor with $\overline{\overline{\epsilon}} = \overline{\overline{\epsilon}}^{+}$, which contradicts (17$a$). Consider a slab region filled with the plasma with magnetic field $\overline{B}$ perpendicular to the slab [Fig. 5.5.4]. Assume that, when a linearly polarized plane wave is transmitted through the slab, the Faraday rotation causes the field vector to rotate 45° in the increasing ϕ direction. Let a current sheet with $\overline{J}_a$ on one side of the slab produce a plane wave polarized in the direction $\phi = 0°$, and a current sheet with $\overline{J}_b$ on the other side of the slab produce a plane wave polarized in the direction $\phi = 45°$. Let $\overline{J}_a$ be source a and $\overline{J}_b$ be source b. The reaction of $< a, b >$ is seen to be zero because the plane wave, as produced by $\overline{J}_b$, is polarized with $\overline{E}_b$ perpendicular to $\overline{J}_a$ after transmitting through the slab, while the reaction of $< b, a >$ is nonzero because $\overline{E}_a$ and $\overline{J}_b$ are in the same direction. Thus $< a, b > \neq < b, a >$ and the Faraday rotation effect is nonreciprocal.

The optical activity also rotates polarization vectors, but the effect is reciprocal. For instance, a quartz crystal exhibits optical rotatory power and can be described as a bianisotropic medium with constitutive relations satisfying (17). Let the slab region in Figure 5.5.4 be filled with an optically active medium such as quartz. The electric field vector will be rotated 45° in increasing ϕ when transmitted upward and rotated 45° in decreasing ϕ when transmitted downward. Thus we have the reaction $< a, b > = < b, a >$.

The reciprocity theorem can be extended as follows. With respect to source a, we write

$$-\nabla \times \overline{E}_a = -i\omega(\overline{\overline{\mu}} \cdot \overline{H} + \overline{\overline{\zeta}} \cdot \overline{E}) + \overline{M}_a \tag{18a}$$

$$\nabla \times \overline{H}_a = -i\omega(\overline{\overline{\epsilon}} \cdot \overline{E} + \overline{\overline{\xi}} \cdot \overline{H}) + \overline{J}_b \tag{18b}$$

The medium is characterized by $\overline{\overline{\mu}}$, $\overline{\overline{\epsilon}}$, $\overline{\overline{\xi}}$, and $\overline{\overline{\zeta}}$. With respect to source b, we use constitutive relations characterized by $\overline{\overline{\mu}}^{C}, \overline{\overline{\epsilon}}^{C}, \overline{\overline{\xi}}^{C}$, and $\overline{\overline{\zeta}}^{C}$ such that

$$\overline{\overline{\mu}}^{C} = \overline{\overline{\mu}}^{T} \tag{19a}$$

$$\overline{\overline{\epsilon}}^{C} = \overline{\overline{\epsilon}}^{T} \tag{19b}$$

$$\overline{\overline{\xi}}^{C} = -\overline{\overline{\zeta}}^{T} \tag{19c}$$

$$\overline{\overline{\zeta}}^{C} = -\overline{\overline{\xi}}^{T} \tag{19d}$$

and call this medium the complementary medium. Maxwell's equations for source b in the complementary medium become

$$-\nabla \times \overline{E}_b^C = -i\omega(\overline{\overline{\mu}}^C \cdot \overline{H}_b^C + \overline{\overline{\zeta}}^C \cdot \overline{E}_b^C) + \overline{M}_b \tag{20a}$$

$$\nabla \times \overline{H}_b^C = -i\omega(\overline{\overline{\epsilon}}^C \cdot \overline{E}_b^C + \overline{\overline{\xi}}^C \cdot \overline{H}_b^C) + \overline{J}_b \tag{20b}$$

where $\overline{E}_b^C$ and $\overline{H}_b^C$ denote the fields produced by $\overline{J}_b$ and $\overline{M}_b$ in the complementary medium. If we define a new reaction

$$< b, a >^C = \iiint_V dV\,(\overline{J}_b \cdot \overline{E}_a^C - \overline{M}_b \cdot \overline{H}_a^C)$$

we find from (18) and (20) that

$$< a, b > = < b, a >^C \tag{21}$$

This result may be called the modified reciprocity theorem, which states that the reaction $< a, b >$ of source a caused by source b in a bianisotropic medium is equal to the reaction $< b, a >^C$ of source b caused by source a in the complementary medium. The medium is reciprocal if the complementary medium is identical to the original medium.

5.6 Stationary Formulas and Rayleigh-Ritz Procedure

Consider a cavity at resonance. We want to calculate the resonant frequency, but we do not know the precise field distribution inside the cavity. Nevertheless, we can find a formula that expresses resonant frequencies in terms of field distributions. We can then assume a field distribution and calculate the resonant frequency in terms of the assumed field. If the formula is stationary, we can come closer to the true resonant frequency than is possible by using a nonstationary formula. Consider the formula $y = f(x)$. We want to calculate y at $x = x_0$, but we do not know the precise value of x_0. We assume that $x = x_0 + p$, where p is a parameter characterizing the deviation from x_0. The formula $y = f(x) = f(x_0 + p)$ can be expanded around x_0 in a Taylor series. We call the formula stationary at $p = 0$ if

$$\left.\frac{\partial f}{\partial p}\right|_{p=0} = 0 \tag{1}$$

When this condition is satisfied, $f(x) = f(x_0) + (1/2)p^2 f^{(2)}(x_0) + \cdots$ and the deviation of $f(x)$ from $f(x_0)$ is of order p^2. Clearly, the stationary formula has an extremum at $p = 0$. When p is complex, we have a saddle point at $p = 0$.

a. Stationary Formula for Resonator Wavenumbers

We shall derive a stationary formula for the resonant frequency of a cavity with an assumed electric field. It is appropriate to mention that a stationary formula involving an assumed magnetic field or a mixture of electric and magnetic fields can be similarly derived. In terms of the assumed electric field (with subscript a), Maxwell's equation gives

$$\begin{aligned} \overline{J}_a &= i\omega\epsilon\overline{E}_a + \nabla \times \overline{H}_a \\ &= \frac{1}{i\omega\mu}\left[-k^2\overline{E}_a + \nabla \times (\nabla \times \overline{E}_a)\right] \end{aligned} \tag{2}$$

inside the cavity. On the cavity wall, we have

$$\overline{M}_s = \hat{n} \times \overline{E}_a \tag{3}$$

where $\hat{n}$ is the unit vector normal to the cavity wall and directed outwards. This magnetic surface current sheet will be zero if the $\overline{E}$ field is the exact electric field, because of the boundary conditions. Since the electric field will be assumed, there is no guarantee that it satisfies the boundary conditions.

Using the definition for reaction, we form a reaction for the cavity,

$$\begin{aligned} < a, a > &= \iiint_V dV\, \overline{J}_a \cdot \overline{E}_a - \oiint_S dS\, \overline{M}_s \cdot \overline{H}_a \\ &= \frac{1}{i\omega\mu}\left\{-k^2 \iiint_V dV\, \overline{E}_a^2 + \iiint_V dV\, (\nabla \times \overline{E}_a)^2 \right. \\ &\qquad \left. + 2\oiint_S dS\, \hat{n} \cdot \left[(\nabla \times \overline{E}_a) \times \overline{E}_a\right]\right\} \end{aligned} \tag{4}$$

In the derivation, we made use of the identities

$$\hat{n} \times \overline{E}_a \cdot \nabla \times \overline{E}_a = -\hat{n} \cdot \left[(\nabla \times \overline{E}_a) \times \overline{E}_a\right]$$

and

$$\overline{E}_a \cdot \nabla \times (\nabla \times \overline{E}_a) = (\nabla \times \overline{E}_a)^2 + \nabla \cdot \left[(\nabla \times \overline{E}_a) \times \overline{E}_a\right]$$

We require that reaction $< a,a >$ be equal to the true reaction of the cavity $< c,c >$, with c standing for "correct". What is the true reaction of the cavity? Inside the cavity, where the field is nonzero, the source is zero. On the cavity walls, where the source is nonzero, the field is zero. Thus $< c,c >= 0$. When $< a,a >$ is set equal to $< c,c >$ equal to zero, we obtain from (4) a formula for the resonant wavenumber k^2 :

$$k^2 = \frac{\iiint_V dV \left(\nabla \times \overline{E}_a\right)^2 + 2 \oiint_S dS\, \hat{n} \cdot \left(\nabla \times \overline{E}_a\right) \times \overline{E}_a}{\iiint_V dV\, \overline{E}_a \cdot \overline{E}_a} \tag{5}$$

Note that this formula is exact if $\overline{E}_a$ is the exact field. We would like to find out whether this formula is stationary. Let $\overline{E}_a = \overline{E} + p\overline{e}$ and denote (5) by $k^2 = N(p)/D(p)$. We wish to examine

$$\left.\frac{\partial k^2}{\partial p}\right|_{p=0} = \frac{D(0)N'(0) - N(0)D'(0)}{D^2(0)} = \frac{N'(0) - k^2 D'(0)}{D(0)}$$

since $N(0) = k^2 D(0)$. Differentiating the numerator of (5) and setting $p = 0$, we obtain

$$\begin{aligned} N'(0) =& 2 \iiint_V dV \left(\nabla \times \overline{E}\right) \cdot \left(\nabla \times \overline{e}\right) + 2 \oiint_S dS\, \hat{n} \cdot \left(\nabla \times \overline{E}\right) \times \overline{e} \\ =& 2k^2 \iiint_V dV\, \overline{e} \cdot \overline{E} = k^2 D'(0) \end{aligned}$$

Here we used the fact that $\hat{n} \times \overline{E} = 0$ on the boundary surface S, the identity $(\nabla \times \overline{E}) \cdot (\nabla \times \overline{e}) = \nabla \cdot (\overline{e} \times \nabla \times \overline{E}) + \overline{e} \cdot \nabla \times \nabla \times \overline{E}$, and the wave equation $\nabla \times \nabla \times \overline{E} = k^2 \overline{E}$. From these results we have proved that $\partial k^2/\partial p|_{p=0} = 0$. Therefore, by requiring that the reaction caused by an assumed field be the same as the reaction attributable to the true field, we obtained a stationary formula for the resonant frequency of cavity.

Consider a circular cavity as shown in Figure 5.6.1. The exact field is known to be, for the fundamental mode, $\overline{E} = \hat{z} E_0 J_0(k\rho)$. The exact resonant wavenumber, which is the first root of $J_0(ka)$,

$$ka = 2.405 \qquad \text{or} \qquad k^2 a^2 = 5.784$$

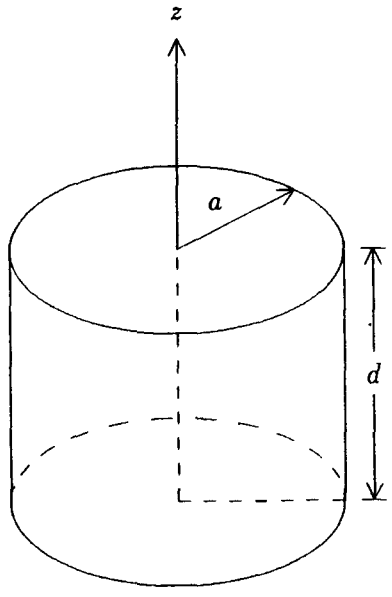


Figure 5.6.1 Circular cavity resonator.

is also known. We now use (5) to estimate the resonant wavenumber. Let us assume a trial field

$$\overline{E}_a = \hat{z} \cos \frac{\pi \rho}{2a}$$

This trial field satisfies the boundary condition at $\rho = a$, and it is not a solution to the wave equation. The curl of $\overline{E}_a$ can be calculated:

$$\nabla \times \overline{E}_a = \hat{\phi} \frac{\pi}{2a} \sin \frac{\pi \rho}{2a}$$

Stationary formula (5) becomes

$$k^2 a^2 = \frac{\int_0^d dz \int_0^{2\pi} d\phi \int_0^a \rho d\rho \, (\pi/2)^2 \sin^2(\pi\rho/2a)}{\int_0^d dz \int_0^{2\pi} d\phi \int_0^a \rho d\rho \cos^2(\pi\rho/2a)} = 5.830$$

This is very close to the exact solution.

Next, we assume the trial field

$$\overline{E}_a = \hat{z} \left(1 + A \frac{\rho}{a}\right) \tag{6}$$

where A is a constant. We shall determine the constant A by using the Ritz procedure as illustrated below. Applying (5), we obtain

$$k^2a^2 = \frac{\int_0^d dz \int_0^{2\pi} d\phi \int_0^a \rho d\rho\, A^2 - \int_0^d dz \int_0^{2\pi} d\phi\, 2a^2 A(1+A)}{\int_0^d dz \int_0^{2\pi} d\phi \int_0^a \rho d\rho\, [1+A(\rho/a)]^2}$$

$$= \frac{-2A - (3/2)A^2}{(1/2) + (2/3)A + (1/4)A^2} \tag{7}$$

requiring k^2a^2 to be stationary with respect to A. If $\partial(k^2a^2)/\partial A = 0$, we find $A = -1$ or -2. Inserting $A = -1$ into (7) yields $k^2a^2 = 6$. The value of $A = -1$ also enables the trial field to satisfy the boundary condition at $\rho = a$. The other value of $A = -2$ gives rise to a negative k^2a^2 and is discarded.

The Ritz procedure can then be extended to n parameters A_l, $l = 1, 2, 3, \ldots, n$, which are then determined from the n equations $\partial k^2/\partial A_l = 0$, $l = 1, 2, \ldots n$. As an example, we assume a trial field characterized by parameters A_1 and A_2 and write

$$\overline{E} = \hat{z}(1 + A_1\frac{\rho}{a} + A_2\frac{\rho^2}{a^2})$$

Then

$$\nabla \times \overline{E} = -\hat{\phi}(A_1 + 2A_2\frac{\rho}{a})/a$$

Inserting in the stationary formula and performing the integration, we obtain

$$k^2a^2 = -\frac{10[18A_2^2 + (28A_1 + 24)A_2 + 9A_1^2 + 12A_1]}{10A_2^2 + (24A_1 + 30)A_2 + 15A_1^2 + 40A_1 + 30}$$

The parameters A_1 and A_2 are determined by requiring that

$$\frac{\partial(k^2a^2)}{\partial A_1} = 0, \qquad \frac{\partial(k^2a^2)}{\partial A_2} = 0$$

These yield two equations:

$$38A_2^3 + (90A_1 + 84)A_2^2 + (51A_1^2 + 45A_1 - 60)A_2 - 45A_1^2 - 135A_1 - 90 = 0$$

$$(76A_1 + 150)A_2^2 + (180A_1^2 + 600A_1 + 540)A_2 + 102A_1^3 + 461A_1^2 + 720A_1 + 360 = 0$$

Solving these equations gives $A_1 = -0.7817$ and $A_2 = -0.1834$. With these values for A_1 and A_2, we obtain

$$k^2a^2 = 5.934$$

This value is closer to the exact solution than the one we would have obtained by using the trial field (6). With these values of A_1 and A_2, we see that the trial field satisfies neither the wave equation nor the boundary condition. Solutions from other values of A_1 and A_2 are not calculated and may correspond to higher resonant wavenumbers. As a final remark, we note that in using the Ritz procedure it is often advisable to choose the trial field components from an orthogonal complete set of functions.

b. Stationary Formula for Antenna Impedance

We now derive a stationary formula for antenna self-impedance. Consider an antenna made of perfect conductor excited by a current source I. The self-reaction of this antenna is due entirely to the current at the terminal. The reason is that on the conducting surface, where the surface current is not zero, the field is zero. The self-reaction is equal to $-VI$ at the terminal. To calculate the self-impedance, we maintain the same terminal current I and assume a current distribution on the surface of the antenna. The self-reaction $< a, a >$ is then calculated. We require that this reaction be approximately equal to the correct reaction $< c, c >$:

$$< a, a > \approx < c, c > = -VI \tag{8}$$

The input impedance is approximated by

$$Z_{in} = -\frac{< c, c >}{I^2} \approx -\frac{< a, a >}{I^2} \tag{9}$$

Let $a = c + pe$. We see that

$$\frac{\partial Z_{in}}{\partial p} = -\frac{1}{I^2}(< e, c > + < c, e >) = -\frac{2}{I^2} < e, c > = 0$$

The second equality follows from reciprocity. The last equality is due to the fact that at the terminal, where the field is not zero, the error

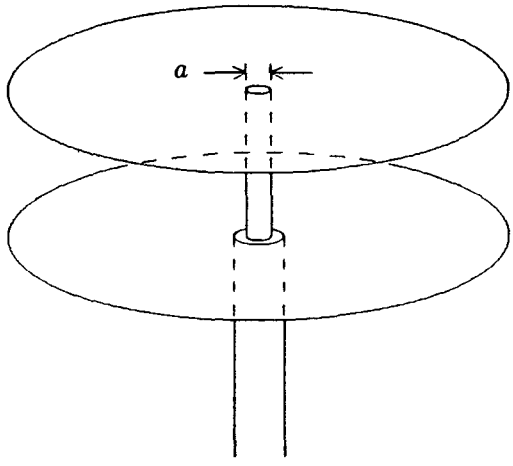


Figure 5.6.2 Probe excitation of a radial parallel-plate waveguide.

source e is zero; everywhere else the correct field c is zero. Thus the formula for the input impedance is proved to be stationary.

We apply the stationary formula for input impedance to a radial parallel-plate waveguide excited by a probe of diameter a [Fig. 5.6.2]. The source terminal of the probe is outside the waveguide. For the TEM mode, we assume that the trial current is uniform along the probe, $\overline{J}_a = \hat{z}I/\pi a$. The electric field generated by this current is

$$\overline{E}_a = \hat{z}\left[-\frac{k^2 I}{4\omega\epsilon} H_0^{(1)}(k\rho)\right]$$

Letting d denote the distance between the two plates, we find the input impedance to be

$$Z_{in} = \frac{\eta}{4} kd\, H_0^{(1)}(ka/2)$$

by using the stationary formula (9).

c. Stationary Formula for Scattering

We shall now derive a stationary formula for scattering problems. Consider a transmitter, a receiver, and a conducting scatterer [Fig. 5.6.3]. The receiver receives a wave composed of two components; one directly from the transmitter in the absence of the scatterer, and one

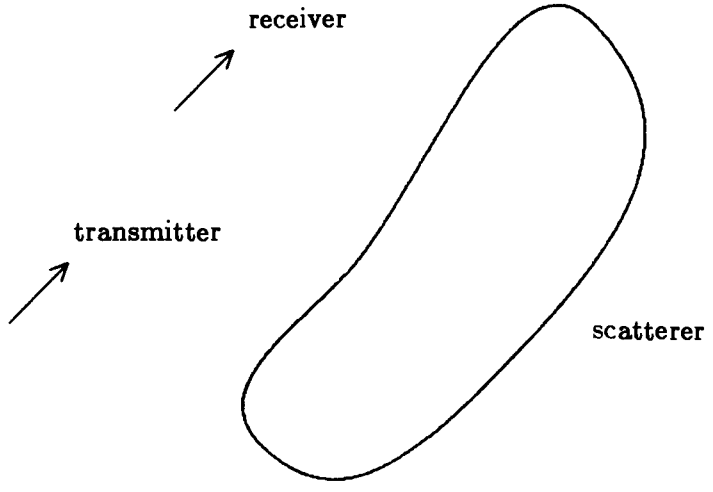


Figure 5.6.3 A transmitter, a receiver, and a conducting scatterer.

originating from currents on the scatterer induced by the transmitter alone. Note that the receiver also induces currents on the scatterer. We want to find a stationary formula for the scattered field V_s as received by the receiver, $V_s = - < i, t >$, where i denotes the receiver source current, and t the field produced by the transmitter-induced currents on the scatterer. We anticipate that the stationary formula will comprise of surface integrals over the scatterer surface. By reciprocity, $V_s = - < i, c_t > = - < c_t, i >$, where, in $< c_t, i >$, c_t denotes the current induced by the transmitter on the scatterer, and i the field produced by the receiver in the absence of the scatterer. In the presence of the scatterer, the field on the scatterer is equal and opposite to i because the scatterer is a conducting body. The field on the scatterer combined with the field in the absence of the scatterer gives a zero field at the scatterer. Let this field be denoted by c_r. We then have on the scatterer surface

$$V_s = - < c_t, i > = < c_t, c_r > \tag{10}$$

Thus far, we have established that the signal received by the receiver is equal to the reaction between the current induced on the scatterer by the transmitter and the field generated on the scatterer by the receiver current.

The letter c in (10) stands for "correct". To calculate V_s, we approximate $< c_t, c_r >$ and write

$$V_s = < c_t, c_r > \approx < a_t, a_r > \tag{11}$$

where a stands for "assumed". The formula is a stationary formula provided that the following constraints are met:

$$< a_t, a_r > = < c_t, a_r > = < a_t, c_r > \tag{12}$$

where c_t is the correct current induced on the scatterer by the transmitter, and c_r is the correct field generated on the scatterer because of the receiver current.

To show that the constraints of (12) can lead us to a stationary formula, let us make a general proposition: a reaction $< a, b >$ is stationary if it satisfies the constraints

$$< a, b > = < c_a, b > = < a, c_b > \tag{13}$$

To prove this theorem, let

$$a = c_a + p_a e_a$$
$$b = c_b + p_b e_b$$

where c stands for "correct", and e for "error". By definition, $< a, b >$ is stationary if

$$\left.\frac{\partial < a, b >}{\partial p_a}\right|_{p_a = p_b = 0} = \left.\frac{\partial < a, b >}{\partial p_b}\right|_{p_a = p_b = 0} = 0 \tag{14}$$

which is seen to be true as the constraint $< a, b > = < c_a, b >$ implies $< e_a, b > = 0$, and the constraint $< a, b > = < a, c_b >$ implies $< a, e_b > = 0$.

It is important to note that constraints (13) are sufficient conditions for a formula to be stationary; they are not necessary conditions. Recall that we did not use these constraints in establishing stationary formulas for the resonant wavenumber of a cavity or for the input impedance of an antenna. In fact, the constraints may be violated. For instance, consider the input impedance of an antenna. Here the constraint $< a, a > = < a, c >$ is violated because, on the antenna surface, the correct fields are zero, but the assumed fields are not.

d. Method of Moments

The method of moments is a numerical technique useful in the solution of electromagnetic wave scattering and radiation problems. Consider a perfectly conducting scattering body that, when illuminated with an incident field $\overline{E}_i$ and $\overline{H}_i$, produces scattered fields $\overline{E}_s$ and $\overline{H}_s$. The scattered fields $\overline{E}_s$ and $\overline{H}_s$ may be attributed to surface currents induced on the scattering body by the incident field. Letting the surface currents be $\overline{J}_s$, we have

$$\overline{E}_s(\overline{r}) = i\omega\mu \iint dS' \, \overline{\overline{G}}(\overline{r}, \overline{r}') \cdot \overline{J}_s(\overline{r}') \tag{15}$$

integrating over the surface of the scatterer S. The unknown surface current may be expanded in terms of basis (or expansion) functions $\overline{F}_n(\overline{r}')$ such that

$$\overline{J}_s(\overline{r}') = \sum_n I_n \overline{F}_n(\overline{r}') \tag{16}$$

Once the expansion coefficients I_n are known, the unknown surface current $\overline{J}_s(\overline{r}')$ will be determined and used to find the desired scattered fields.

The method of moments calls for the use of a set of testing functions $\overline{J}_m(\overline{r}')$ with which we form

$$\begin{aligned} \iint_{S_m} dS_m \, \overline{J}_m(\overline{r}') \cdot \overline{E}_s(\overline{r}') &= \sum_n \left[\iint_{S_m} dS_m \, \overline{J}_m(\overline{r}') \cdot i\omega\mu \iint dS' \overline{\overline{G}} \cdot \overline{F}_n \right] I_n \\ &= \sum_n Z_{mn} I_n \end{aligned} \tag{17}$$

where

$$Z_{mn} = \iint_{S_m} dS_m \, \overline{J}_m(\overline{r}') \cdot \overline{F}_n(\overline{r}') \tag{18}$$

and S_m is part of the surface of the perfectly conducting scatterer. The set of testing functions covers the whole surface S of the scattering body.

On the conducting surface, the boundary condition requires that the total tangential electric field $\overline{E}_s + \overline{E}_i = 0$. From (17), we find

$$\iint_{S_m} dS_m \, \overline{J}_m(\overline{r}') \cdot \overline{E}_s(\overline{r}') = - \iint_{S_m} dS_m \, \overline{J}_m(\overline{r}') \cdot \overline{E}_i(\overline{r}') \tag{19}$$

This should be compared with (5.6.10) where $< c_t, c_r > = - < c_t, i >$ although c_t is now actually a_t. Defining

$$V_m = - \iint_{S_m} dS_m \, \overline{J}_m(\overline{r}') \cdot \overline{E}_i(\overline{r}') \tag{20}$$

We find from (17)–(19)

$$Z_{mn} I_n = V_m \tag{21}$$

and consequently the unknown expansion coefficients for the surface currents

$$I_n = Z_{mn}^{-1} V_m \tag{22}$$

Thus the essence of the method of moments lies in the choice of the basis and testing functions, and the numerical inversion of the Z-matrix Z_{mn}. In the point matching technique, we choose the testing functions to be delta functions such that $\overline{J}_m(\overline{r}') = \overline{J}_m \delta(\overline{r}' - \overline{r}_m)$ and cover the scattering body with enough points to ensure a convergent result. On those chosen points, the boundary condition of zero tangential $\overline{E}$ field is enforced. When the same function is used for both basis expansion and testing, the technique is called Galerkin's method.

5.7 Geometrical Optics Limit

Geometrical optics can be treated as a limiting case of very high frequency approximation in Maxwell's theory. Consider the solution

$$\overline{E}(\overline{r}) = \overline{E} \, e^{ik_0 L(\overline{r})} \tag{1}$$

$$\overline{H}(\overline{r}) = \overline{H} \, e^{ik_0 L(\overline{r})} \tag{2}$$

where $k_0 = \omega/c$ and we shall let $k_0 \to \infty$. Substituting in source-free Maxwell's equations and making use of the vector identities $\nabla \times (\overline{A}\phi) = \phi \nabla \times \overline{A} + \nabla\phi \times \overline{A}$, and $\nabla \cdot (\overline{A}\phi) = \nabla\phi \cdot \overline{A} + \phi \nabla \cdot \overline{A}$, we find that in isotropic media,

$$\nabla L(\overline{r}) \times \overline{H} + \frac{n}{\eta} \overline{E} = \frac{i}{k_0} \nabla \times \overline{H} \tag{3}$$

$$\nabla L(\overline{r}) \times \overline{E} - n\eta \overline{H} = \frac{i}{k_0} \nabla \times \overline{E} \tag{4}$$

$$\nabla L(\overline{r}) \cdot \overline{E} = \frac{i}{k_0} \nabla \cdot \overline{E} \tag{5}$$

$$\nabla L(\overline{r}) \cdot \overline{H} = \frac{i}{k_0} \nabla \cdot \overline{H} \tag{6}$$

where $n = c\sqrt{\mu\epsilon}$ is the refractive index and $\eta = \sqrt{\mu/\epsilon}$ is the characteristic impedance for the isotropic media.

In the high-frequency limit we omit the right-hand sides of (3)–(6) and obtain the governing equations for geometrical optics

$$\nabla L \times \overline{H} + \frac{n}{\eta}\overline{E} = 0 \tag{7}$$

$$\nabla L \times \overline{E} - n\eta\overline{H} = 0 \tag{8}$$

$$\nabla L \cdot \overline{E} = 0 \tag{9}$$

$$\nabla L \cdot \overline{H} = 0 \tag{10}$$

This set of equations is now independent of frequency.

Substituting (8) in (7) and making use of (9) we obtain the eikonal equation in geometrical optics

$$|\nabla L(\bar{r})|^2 = n^2 \tag{11}$$

The phase function $L(\bar{r})$ is also called the eikonal. The geometrical wavefronts are described by $L(\bar{r}) = \textit{constant}$. Let the unit normal to a wavefront be $\hat{s}$. We see from (11) that

$$\nabla L = \hat{s}n \tag{12}$$

where $\hat{s}n$ is called the ray vector.

We see from (9) and (10) that the electric and magnetic fields are perpendicular to the normal of the wavefront. The time-average Poynting's vector is

$$\begin{aligned} < \overline{S} > &= \frac{1}{2} Re\left\{\overline{E} \times \overline{H}^*\right\} = \frac{1}{2n\eta} Re\left\{\overline{E} \times (\nabla L \times \overline{E})^*\right\} \\ &= \frac{1}{2n\eta}(\overline{E} \cdot \overline{E}^*)\nabla L = \frac{2c}{n^2} < W_e > \nabla L = \hat{s}\frac{2c}{n} < W_e > \end{aligned} \tag{13}$$

or, in terms of $\overline{H}$,

$$\begin{aligned} < \overline{S} > &= -\frac{\eta}{2n} Re\left\{(\nabla L \times \overline{H}) \times \overline{H}^*\right\} \\ &= \frac{\eta}{2n}(\overline{H} \cdot \overline{H}^*)\nabla L = \frac{2c}{n^2} < W_m > \nabla L = \hat{s}\frac{2c}{n} < W_m > \end{aligned} \tag{14}$$

where $< W_e > = \epsilon\overline{E} \cdot \overline{E}^*/4$ is the time-average electric energy, and $< W_m > = \mu\overline{H} \cdot \overline{H}^*/4$ is the time-average magnetic energy. We find

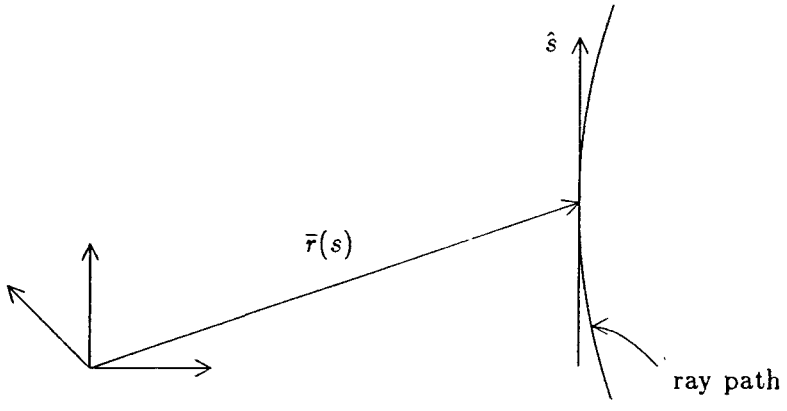


Figure 5.7.1 Ray path in geometrical optics.

by equating (13) and (14) that $< W_e > = < W_m > = < W > /2$ where $< W >$ is the total time-average stored energy. Thus

$$< \overline{S} > = \hat{s}\frac{c}{n} < W > = \hat{s}v < W > \tag{15}$$

where $v = c/n$ is the velocity of the electromagnetic wave in the medium. We conclude that the time-average Poynting's power vector is in the direction of the ray vector $\hat{s}n$ and its magnitude is equal to the product of the electromagnetic energy $< W >$ and the velocity v.

The intensity I is defined as the absolute value of the time-average Poynting vector

$$I = |< \overline{S} >| = v < W >$$

Poynting's theorem states that

$$\nabla \cdot \overline{S} = -i\omega(\epsilon|E|^2 - \mu|H|^2)$$

Taking time-average of the above equation we obtain

$$\nabla \cdot (\hat{s}I) = 0$$

This is the conservation equation for the intensity I.

Let the position vector along a ray path [Fig. 5.7.1] be denoted by $\bar{r}(s)$, expressed in terms of the arc length s as the parameter for the path. Since $d\bar{r}/ds = \hat{s}$, we find from (12)

$$\frac{d^2\bar{r}}{ds^2} = \frac{d\bar{r}}{ds} \cdot \nabla\left(\frac{d\bar{r}}{ds}\right) = \hat{s} \cdot \nabla(\hat{s}) = \frac{\nabla L}{n} \cdot \nabla\left(\frac{\nabla L}{n}\right)$$

For homogeneous media with n being a constant, it is seen that

$$\frac{d^2\bar{r}}{ds^2} = \frac{1}{2}\nabla\left(\frac{|\nabla L|^2}{n^2}\right) = 0$$

in view of (11). Thus the ray path is a straight line. When n is spatially dependent, we make use of the vector identity

$$\nabla(\overline{A}\cdot\overline{B}) = (\overline{A}\cdot\nabla)\overline{B} + (\overline{B}\cdot\nabla)\overline{A} + \overline{A}\times(\nabla\times\overline{B}) + \overline{B}\times(\nabla\times\overline{A})$$

and find that

$$\begin{aligned}\frac{d^2\bar{r}}{ds^2} &= \frac{1}{2}\nabla\left(\frac{|\nabla L|^2}{n^2}\right) - \frac{\nabla L}{n}\times\left(\nabla\frac{1}{n}\times\nabla L\right)\\ &= \frac{\nabla L}{n}\left(\nabla L\cdot\nabla\frac{1}{n}\right) - \left(\nabla\frac{1}{n}\right)\frac{|\nabla L|^2}{n} + \frac{1}{2}\nabla\left(\frac{|\nabla L|^2}{n^2}\right)\\ &= \frac{\nabla L}{n}\left(\nabla L\cdot\nabla\frac{1}{n}\right) - n\nabla\frac{1}{n}\end{aligned}$$

Thus the ray is curved in inhomogeneous media.

Taking the curl of (12), we also have

$$\nabla\times(\hat{s}n) = 0 \tag{16}$$

We now derive Snell's law by integrating (16) around a ribbon-like contour across the boundary separating two media with refractive indices n_1 and n_2 [Fig. 5.7.2]. Applying Stokes' theorem and letting the ribbon width $\delta\to 0$, we have

$$\iint d\overline{S}\cdot\nabla\times(\hat{s}n) = \oint_C d\bar{l}\cdot\hat{s}n = 0 \tag{17}$$

where $d\overline{S}$ is the unit vector perpendicular to the ribbon area and $d\bar{l}$ is the differential line element along the closed contour of the ribbon. The contributions from the two sides perpendicular to the surface are neglected because they are proportional to δ. Thus the tangential components of the ray vectors are continuous across the boundary. For the transmitted ray we find from (17)

$$n_1\sin\theta_i = n_2\sin\theta_t \tag{18}$$

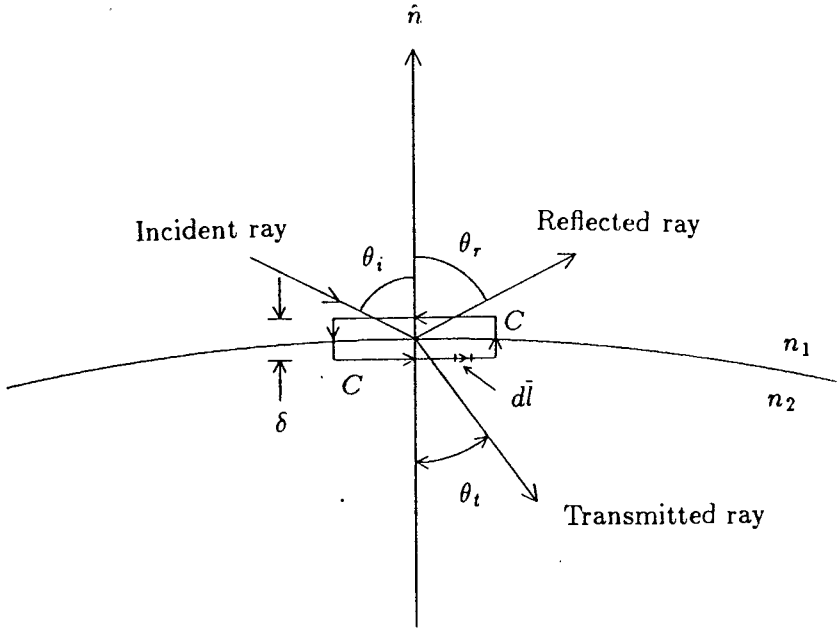


Figure 5.7.2 Derivation of Snell's law.

For the reflected ray, we use $\theta_t = \pi - \theta_r$ and $n_2 = n_1$. From (18), we find $n_1 \sin\theta_i = n_1 \sin\theta_r$. Thus the angle of reflection θ_r is equal to the angle of incidence θ_i. Note that Snell's law has also been derived from phase matching wave vectors for plane waves. The above derivation is valid, provided that the radii of curvature of the incident wave and of the boundary surface are large compared to the wavelength. Due to the curl-free property of the ray vector, we have

$$\oint_C d\bar{r} \cdot \hat{s}n = 0 \tag{19}$$

along any closed path C. This equation is true even across the boundary of two different media by virtue of Snell's law.

Consider the problem of reflection of a wavefront S_1 by a surface S and the reflected rays form another wavefront S_2 [Fig. 5.7.3]. Note that A_1AA_2 and B_1BB_2 are the ray paths in the direction of $\hat{s}$ and A_1B_1 and A_2B_2 are wavefronts perpendicular to the ray paths. Applying

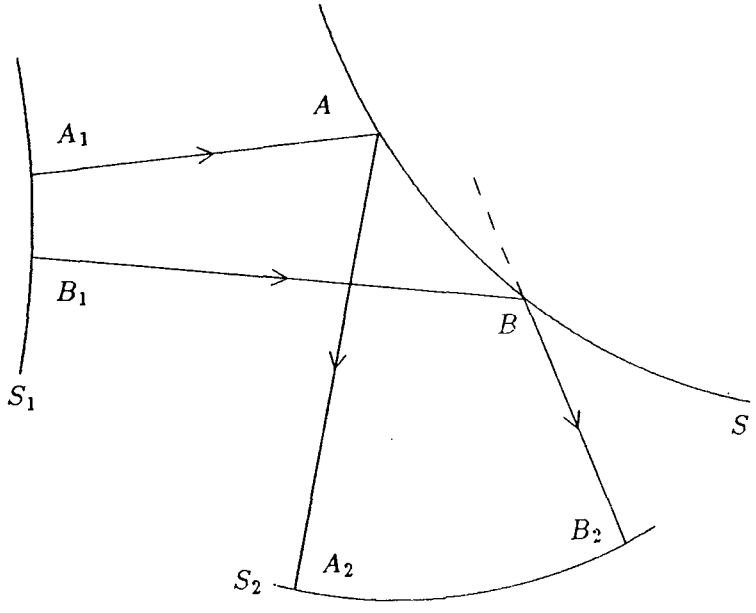


Figure 5.7.3 Reflection of wavefront S_1 by a surface S.

(19), we find

$$\left\{\int_{A_1AA_2} + \int_{A_2B_2} + \int_{B_2BB_1} + \int_{B_1A_1}\right\} d\bar{r} \cdot \hat{s}n = 0$$

The integrals along A_2B_2 and B_1A_1 are zero because $d\bar{r} \cdot \hat{s} = 0$. We have

$$\int_{A_1AA_2} d\bar{r} \cdot \hat{s}n - \int_{B_1BB_2} d\bar{r} \cdot \hat{s}n = 0$$

The vectors $d\bar{r}$ and $\hat{s}n$ are in the same direction on the optical path. The integrals define the optical path lengths along paths A_1AA_2 and B_1BB_2. We conclude

$$\int_{A_1AA_2} n\,ds = \int_{B_1BB_2} n\,ds \tag{20}$$

Thus the optical path length between any two wavefronts is the same for all rays. The above theorem is easily generalized to include cases of refraction.

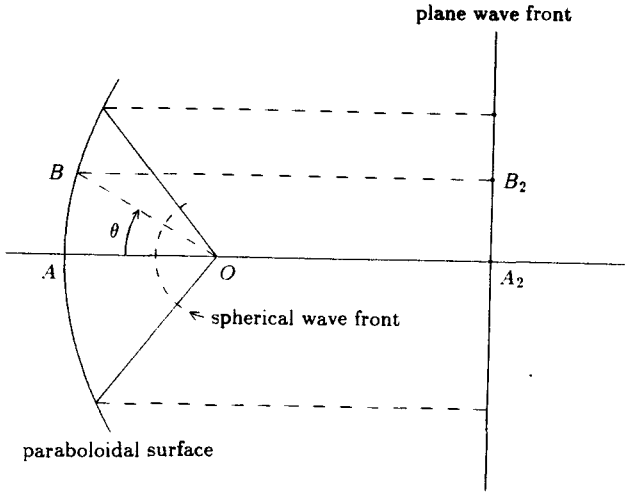


Figure 5.7.4 Converting spherical phase front to plane phase front.

As an example, consider a reflecting surface that converts a spherical phase front to a plane phase front. Consider the ray paths as shown in Figure 5.7.4. In view of (20), we have

$$\overline{OB} + \overline{BB_2} = \overline{OA} + \overline{AA_2} \tag{21}$$

Since all reflected rays are parallel to AOA_2, we have

$$\overline{BB_2} = \overline{OB}\cos\theta + \overline{OA_2} \tag{22}$$

Substituting in (21), we find

$$\begin{aligned} \overline{OB}(1+\cos\theta) &= \overline{OA} - \overline{OA_2} + \overline{AA_2} = 2\overline{OA} \\ r(1+\cos\theta) &= 2f \end{aligned} \tag{23}$$

This equation describes a parabola. Revolving around the axis OA_2 we see that the reflecting surface takes a paraboloidal shape.

A hyperboloidal surface converts one spherical phase front into another spherical phase front. As shown in Figure 5.7.5, we obtain from (20)

$$\overline{OB} + \overline{BB_2} = \overline{OA} + \overline{AA_2} \tag{24}$$

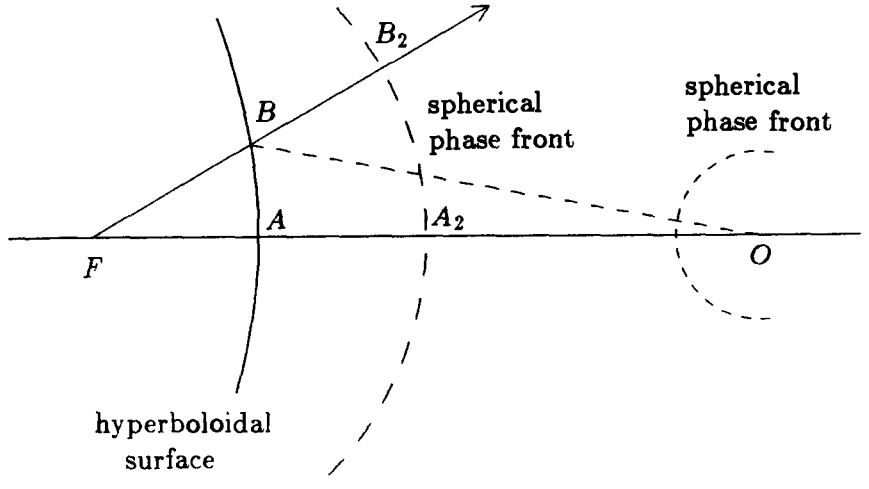


Figure 5.7.5 Converting a spherical phase front to another spherical phase front.

The center of the spherical phase front of the reflected rays is at F, since

$$\overline{BB_2} = \overline{FB_2} - \overline{FB} = \overline{FA_2} - \overline{FB}$$

We find from (24)

$$\overline{OB} - \overline{FB} = \overline{OA} + \overline{AA_2} - \overline{FA_2} = \overline{OA} - \overline{FA} \tag{25}$$

This equation determines a hyperbola. Revolution around axis FO results in a hyperboloidal surface. Reflectors of this type were originally designed by Cassegrain for optical telescopes.

Lens Antennas

We first establish the optical path-length theorem for refraction in geometrical optics. Consider the boundary surface S separating air and a dielectric medium with refractive index n as shown in Figure 5.7.6. The two rays A_1AA_2 and B_1BB_2 emerge from the constant phase front S_1, refracted by the medium and form the new phase front S_2.

The field amplitudes of rays have the dependence $e^{i\omega L(\bar{r})/c}$ and

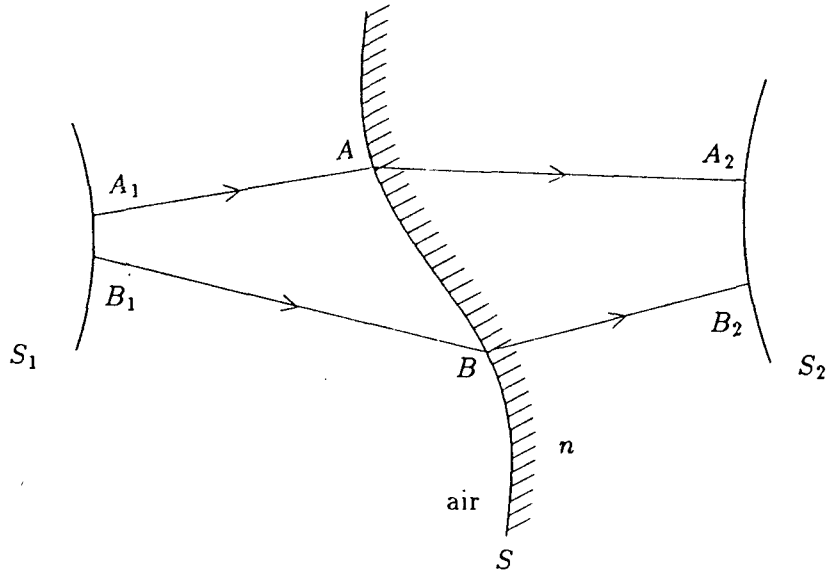


Figure 5.7.6 Establishing optical path-length theorem.

the ray vector $n\hat{s} = \nabla L(\bar{r})$ has the curl-free property

$$\nabla \times n\hat{s} = 0 \tag{26}$$

where n is the refractive index of air or the dielectric depending on the ray location. Stokes' theorem yields

$$\int_{A_1A} d\bar{r} \cdot \hat{s} + \int_{AA_2} d\bar{r} \cdot \hat{s}n + \int_{A_2B_2} d\bar{r} \cdot \hat{s}n + \int_{B_2B} d\bar{r} \cdot \hat{s}n + \int_{BB_1} d\bar{r} \cdot \hat{s} + \int_{B_1A_1} d\bar{r} \cdot \hat{s} = 0$$

By the boundary condition of the ray vector at the surface S, the rays A_1A to AA_2 and similarly B_1B to BB_2 satisfy Snell's law. The integrals along A_1B_1 and A_2B_2 are zero because $d\bar{r} \cdot \hat{s} = 0$. We thus have

$$\int_{A_1A} d\bar{r} \cdot \hat{s} + \int_{AA_2} d\bar{r} \cdot \hat{s}n = \int_{B_1B} d\bar{r} \cdot \hat{s} + \int_{BB_2} d\bar{r} \cdot \hat{s}n \tag{27}$$

or simply

$$A_1A + n\,AA_2 = B_1B + n\,BB_2 \tag{28}$$

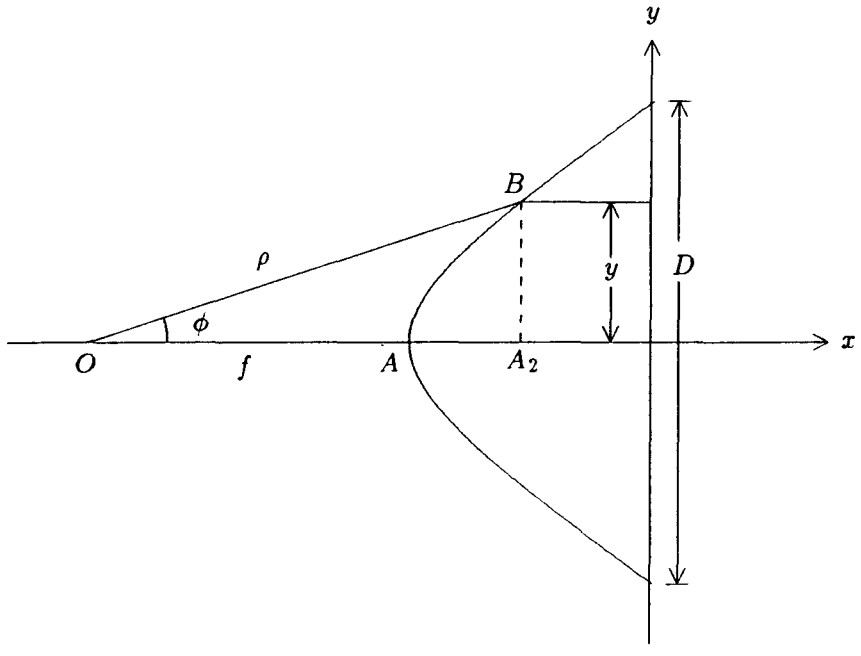


Figure 5.7.7 One-surface lens to convert a spherical phase front to a plane phase front.

This is the optical path length theorem which states that the optical path lengths between two constant phase fronts are equal.

We now construct a one-surface lens that converts a spherical phase front into a plane phase front. By the optical path length theorem, we see from Figure 5.7.7 that

$$\overline{OB} = \overline{OA} + n\,\overline{AA_2} \tag{29}$$

Letting the length from the origin O to the refraction surface be ρ, at an angle ϕ with the x axis, we find $\overline{AA_2} = \rho\cos\phi - f$ where $f = \overline{OA}$. Equation (29) becomes

$$\rho = \frac{(n-1)f}{n\cos\phi - 1} \tag{30}$$

This is the equation for a hyperbola.

Consider a cylindrical lens antenna with the hyperbolic cross-section as described by (30). The antenna is illuminated with a line

source. We shall show that the field amplitude distribution at the aperture with diameter D is tapered. First we note that the total power P_t passing through the strip at y with width dy is

$$P_t = dy\,P(y) \tag{31}$$

where $P(y)$ is power per unit length at y. Neglecting reflections, we see that this total power is equal to that radiated by the line source over the solid angle $d\phi$,

$$P_t = U(\phi)\,d\phi \tag{32}$$

where $U(\phi)$ is power per unit angle from the source. Since $y = \rho \sin\phi$, we obtain from (30)

$$dy = (n-1)f\frac{n-\cos\phi}{(n\cos\phi-1)^2}\,d\phi \tag{33}$$

Equating (31) and (32) and making use of (33), we obtain

$$P(y) = \frac{(n\cos\phi-1)^2}{(n-1)f(n-\cos\phi)}\,U(\phi) \tag{34}$$

In the aperture plane the ratio of the field amplitude $E(y)$ at y to $E(0)$ at $y=0$ corresponding to $\phi = 0$ is equal to the square root of their power ratio,

$$\frac{E(y)}{E(0)} = \sqrt{\frac{P(y)}{P(0)}} = \frac{n\cos\phi - 1}{\sqrt{(n-1)(n-\cos\phi)}} \tag{35}$$

The amplitude ratio becomes zero when $\phi = \cos^{-1}(1/n)$.

The neglect of reflection at the convex surface increases the transmitted power. Of more importance is the reflection from the plane surface which is focused back into the primary feed and causes a mismatch. The reflections at both surfaces may be reduced with a material of small refractive index. The reflection at the plane surface can also be eliminated by a quarter-wavelength plate with refractive index $\sqrt{n}$. The reason follows from transmission line theory where a quarter-wavelength section with characteristic impedance $\sqrt{(Z_0 Z_1)}$ can be shown to match a line with characteristic impedance Z_0 to another

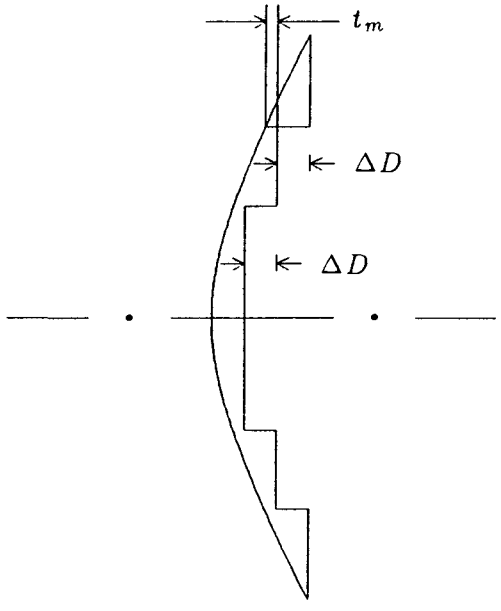


Figure 5.7.8 Zoned dielectric lens.

line with Z_1 as its characteristic impedance. The transmission line theory is applicable here because the wave is normally incident at the plane surface.

To reduce the weight of the lens, zoned dielectric lenses are constructed [Fig. 5.7.8]. From the optical path length theorem (27), we recognize that $d\bar{r} \cdot \hat{s}n = dL$ is the incremental phase change in the field dependence $e^{ik_o L(\bar{r})}$ with $\Delta L(\bar{r}) = \hat{s}n$. Thus we make the zoned step ΔD such that

$$k_o \Delta D\,(n-1) = 2\pi$$

or equivalently

$$\Delta D = \frac{\lambda_o}{n-1}$$

where $\lambda_o = 2\pi/k_o$ is the free-space wavelength. Then the phase difference caused by removal of the dielectric is 2π. When the smallest thickness of the lens is t_m, the maximum thickness of the lens will be $t_m + \lambda_o/(n-1)$.

The conversion of a cylindrical phase front produced by the line

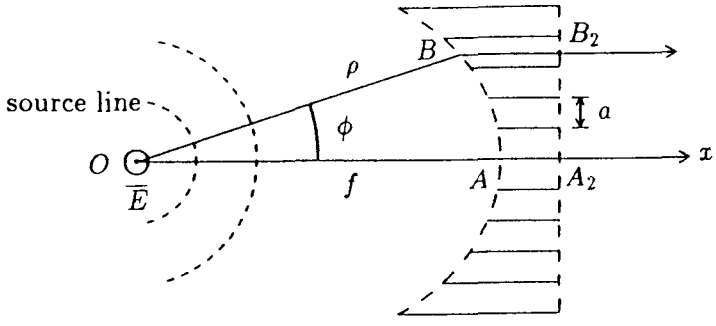


Figure 5.7.9 Metal-plate lens.

source at the focal point to a plane phase front at the aperture plane of the lens can be understood from the point of view of phase retardation due to the dielectric medium. Inside the dielectric the phase velocity is c/n. Close to the x axis, the dielectric is the thickest and the phase retardation is the largest. At the edge of the lens the phase velocity is not slowed down at all. The hyperbolic cross-section provides such distribution in the phase retardation that the resultant phase front is a plane.

The same concept can be used in the understanding of lens antennas made of metal plates. Inside two parallel plates separated by a distance a, the phase velocity of a TE_1 wave is

$$v = c\left[1 - \left(\frac{\pi}{ka}\right)^2\right]^{-\frac{1}{2}} \tag{36}$$

which is seen to be larger than the velocity of light in free space c. The equivalent index of refraction is

$$n = \frac{c}{v} = \sqrt{1 - \left(\frac{\pi}{ka}\right)^2} \tag{37}$$

Note that in order to have a TE_1 wave propagating above cutoff inside the plate waveguide, we must have $\pi < ka$.

Consider a metal-plate lens antenna as shown in Figure 5.7.9. The profile of the plates can be determined by the optical path length theorem with the equivalent refractive index concept. We have

$$\overline{OB} + n\,\overline{BB_2} = \overline{OA} + n\,\overline{AA_2} \tag{38}$$

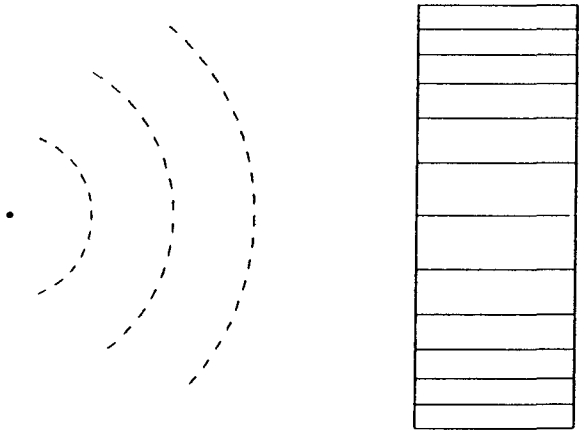


Figure 5.7.10 Metal-plate lens with nonuniform separations.

Letting $\overline{OB} = \rho$ and $\overline{OA} = f$, we find $\overline{BB_2} - \overline{AA_2} = f - \rho\cos\phi$. From (38) we have

$$\rho = \frac{(1-n)f}{1 - n\cos\phi} \tag{39}$$

With $n < 1$, (39) is the equation for an ellipse.

The effect of the longer metal plate waveguides is to speed up the phase-front propagation to form a plane phase front at the aperture plane. The smaller the separation a the larger the phase velocity inside the plate waveguide; thus, metal-plate lens antennas can be made with uniform length but nonuniform separations. Such an arrangement is shown in Figure 5.7.10.

Paraboloidal Reflector Antenna

For the paraboloidal reflector surface as shown in Figure 5.7.11, it is described by the equation

$$r + r\cos\theta = 2f \tag{40}$$

where f is the focal length. Let the plane perpendicular to the z axis at the focal point be called the aperture plane. Equation (40) states that all path lengths measured from the focal point F to the reflector surface and then to the plane surface perpendicular to the axis at the focal point are equal to the constant $2f$.

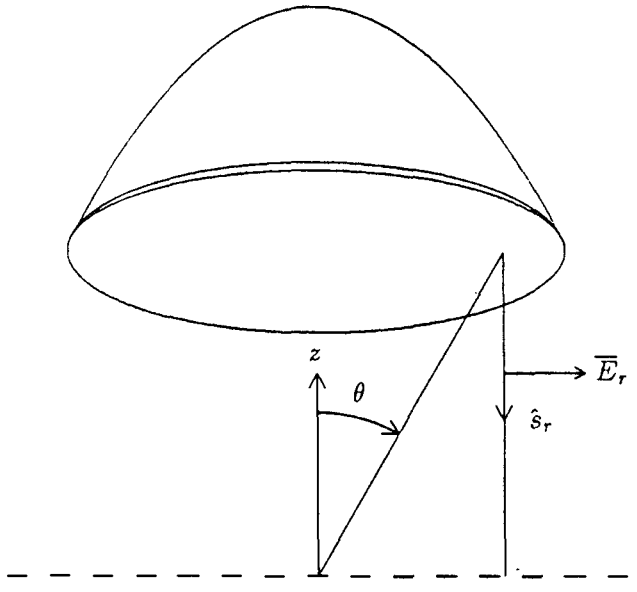


Figure 5.7.11 Paraboloidal reflector.

It is interesting to find by means of geometrical ray optics that the power distribution on the aperture plane is nonuniform. Consider an omnidirectional source at the focal point; the power radiated in the solid angle $d\theta$ rotating about the z axis is proportional to $2\pi \sin\theta \, d\theta$. We write

$$dp = 2\pi C \sin\theta \, d\theta$$

where C is a constant proportional to the source strength. By the conservation law of geometrical optics, this power appears in the aperture plane through a ring of differential width $d\rho$ with differential area $dA = 2\pi\rho \, d\rho$. Thus

$$\frac{dp}{dA} = \frac{C \sin\theta \, d\theta}{\rho \, d\rho} \tag{41}$$

From (40) we have

$$r = f \sec^2 \frac{\theta}{2} \tag{42}$$

and

$$\rho = r \sin\theta = 2f \tan\frac{\theta}{2} \tag{43}$$

Making use of (43), we find (41) becomes

$$\frac{dp}{dA} = \frac{C}{f^2}\cos^4\left(\frac{\theta}{2}\right)$$

On the aperture plane, we have a tapered amplitude distribution. The distribution can be made uniform by making the gain pattern of the feeding source $G(\theta,\phi) = \sec^4(\theta/2)$.

The surface current density at the reflector surface is, under the physical optics approximation,

$$\overline{J}_s = \hat{n} \times (\overline{H}_i + \overline{H}_r) \tag{44}$$

where $\hat{n}$ is the unit normal to the paraboloidal surface, $\overline{H}_i$ is the incident magnetic field, and $\overline{H}_r$ is the reflected magnetic field. On a perfectly conducting surface, $\hat{n} \times \overline{H}_i = \hat{n} \times \overline{H}_r$. In terms of the incident and reflected electric fields,

$$\overline{H}_i = \frac{1}{\eta}\hat{s}_i \times \overline{E}_i \tag{45}$$

$$\overline{H}_r = \frac{1}{\eta}\hat{s}_r \times \overline{E}_r \tag{46}$$

where $\hat{s}_i$ is the unit incident ray vector and $\hat{s}_r$ the unit reflected ray vector. The surface current density becomes

$$\overline{J}_s = \frac{2}{\eta}[\hat{n} \times (\hat{s}_i \times \overline{E}_i)] \tag{47}$$

or

$$\overline{J}_s = \frac{2}{\eta}[\hat{n} \times (\hat{s}_r \times \overline{E}_r)] \tag{48}$$

Now assume that at the feed there is a point source generating an incident field vector

$$\overline{E}_i = \hat{e}_i\, E(\theta,\phi)\frac{e^{i\omega r/c}}{r} \tag{49}$$

on the paraboloidal surface at r where $r \gg \lambda$. The amplitude $E(\theta,\phi)$ is related to the gain factor $G(\theta,\phi)$ of the source by

$$\frac{1}{2\eta}|E(\theta,\phi)|^2 = \frac{P_t}{4\pi}G(\theta,\phi) \tag{50}$$

where P_t is the total power radiated by the source.

We now study the polarization of the reflected field. The electromagnetic boundary condition requires that tangential electric fields vanish at the reflector surface

$$\hat{n} \times (\overline{E}_r + \overline{E}_i) = 0 \tag{51}$$

also

$$\hat{n} \cdot \overline{E}_r = \hat{n} \cdot \overline{E}_i \tag{52}$$

Cross-multiplying (51) by $\hat{n}$ and making use of (52), we find

$$\overline{E}_r = 2\hat{n}(\hat{n} \cdot \overline{E}_i) - \overline{E}_i = \hat{e}_r \, E(\theta, \phi) \frac{e^{i\omega r/c}}{r} \tag{53}$$

The unit vector $\hat{e}_r$ for the reflected wave in (53) is

$$\hat{e}_r = 2\hat{n} \frac{(\hat{n} \cdot \overline{E}_i)}{|\overline{E}_r|} - \frac{\overline{E}_i}{|\overline{E}_r|} = 2\hat{n}(\hat{n} \cdot \hat{e}_i) - \hat{e}_i \tag{54}$$

where we used the fact that $|\overline{E}_r| = |\overline{E}_i|$ as is seen from (53).

We assume that the point source at the focus is linearly polarized in the $\hat{y}$ direction. The unit vector $\hat{e}_i$ is

$$\hat{e}_i = \frac{\hat{r} \times (\hat{y} \times \hat{r})}{|\hat{r} \times (\hat{y} \times \hat{r})|} = \frac{\hat{y} - (\hat{r} \cdot \hat{y})\hat{r}}{|\hat{y} - (\hat{r} \cdot \hat{y})\hat{r}|} \tag{55}$$

Since $\hat{r} = \hat{x} \sin\theta \cos\phi + \hat{y} \sin\theta \sin\phi + \hat{z} \cos\theta$, we find

$$\hat{e}_i = \frac{1}{\sqrt{1 - \sin^2\theta \sin^2\phi}} \{ -\hat{x} \sin^2\theta \sin\phi \cos\theta + \hat{y}(1 - \sin^2\theta \sin^2\phi) - \hat{z} \sin\theta \cos\theta \sin\phi \} \tag{56}$$

in rectangular coordinates.

The unit normal to the reflector $\hat{n}$ can be found by taking the gradient of (42),

$$\nabla \left[f - r \cos^2 \frac{\theta}{2} \right] = \left(-\hat{r} \cos\frac{\theta}{2} + \hat{\theta} \sin\frac{\theta}{2} \right) \cos\frac{\theta}{2}$$

Thus

$$\hat{n} = -\hat{r}\cos\frac{\theta}{2} + \hat{\theta}\sin\frac{\theta}{2} \tag{57}$$

such that $\hat{n}\cdot\hat{n} = 1$. In rectangular coordinates

$$\hat{n} = -\hat{x}\sin\frac{\theta}{2}\cos\phi - \hat{y}\sin\frac{\theta}{2}\sin\phi - \hat{z}\cos\frac{\theta}{2} \tag{58}$$

We now determine the reflected field polarization $\hat{e}_r$. From (55) and (57) we find $\hat{n}\cdot\hat{e}_i = (\hat{\theta}\cdot\hat{y})\sin(\theta/2) = \sin(\theta/2)\cos\theta\sin\phi$. It follows from (54), (56), and (58), that

$$\hat{e}_r = 2\hat{n}(\hat{n}\cdot\hat{e}_i) - \hat{e}_i = \frac{1}{\sqrt{1-\sin^2\theta\sin^2\phi}}\Big\{\hat{x}(1-\cos\theta)\sin\phi\cos\phi + \hat{y}(\cos\theta\sin^2\phi + \cos^2\phi)\Big\} \tag{59}$$

As we had expected, $\hat{e}_r$ has no z component because $\hat{s}_r$ is in the $-\hat{z}$ direction and $\overline{E}_r$ is perpendicular to $\hat{s}_r$. The x component of $\hat{e}_r$ gives the newly generated cross-polarized field at the aperture plane. By geometrical optics, the field at the aperture plane acquires an additional phase $(\omega/c)r\cos\theta$ as compared with $\overline{E}_r$. We have

$$\overline{E}_{ap} = \hat{e}_r E(\theta,\phi)\frac{e^{-i(\omega/c)(r+r\cos\theta)}}{r} \tag{60}$$

For large f/D ratio, the reflector section is relatively flat. We find $\theta \to 0$ and the x component of $\hat{e}_r$, $e_{rx} \to 0$. It is interesting to note that for an electric dipole at the feed, its incident field is proportional to the sine of the angle γ between $\hat{r}$ and the y axis with $\sin\gamma = \sqrt{1-\cos^2\gamma} = \sqrt{1-(\hat{r}\cdot\hat{y})^2} = \sqrt{1-\sin^2\theta\sin^2\gamma}$. This cancels the angular dependence in the denominator of $\hat{e}_r$.

Having determined the surface current distribution on the reflector surface, we now find the radiation fields from

$$\overline{E}(\overline{r}) = i\omega\mu\frac{e^{ikr}}{4\pi r}(\hat{\theta}f_\theta + \hat{\phi}f_\phi)$$

where the vector current moment

$$\overline{f} = \iint dS\,\overline{J}_s(\overline{r}')\,e^{i\overline{k}\cdot\overline{r}'}$$

The surface current density is derived in (48)

$$\overline{J}_s(\overline{r}') = \frac{2}{\eta}\left\{\hat{n} \times [\hat{s}_r \times \overline{E}_r(\overline{r}')]\right\}$$

From (53), the reflected field

$$\overline{E}_r(\overline{r}') = \hat{e}_r E(\theta', \phi') \frac{e^{ikr'}}{r'}$$

Since $\hat{s}_r = -\hat{z}$ and $-\hat{z} \cdot \hat{n} = \cos(\theta/2)$, we find

$$\overline{J}_s(\overline{r}') = \left[\hat{e}_r \cos\frac{\theta'}{2} - \hat{z}(\hat{n} \cdot \hat{e}_r)\right] \frac{2E(\theta', \phi')}{\eta} \frac{e^{ikr'}}{r'} \tag{61}$$

where $\hat{e}_r$ is given by (59) and $\hat{n}$ by (57).

We see that the vector current moment $\overline{f}(\overline{r})$ has a component along $\hat{e}_r$ which is parallel to the $x - y$ plane and a component along the z axis. The z component makes no contribution to E_ϕ because $\hat{\phi}$ is always perpendicular to z. Its contribution to E_θ is vanishingly small around $\theta = 0$ because $\hat{z} \cdot \hat{\theta} = \sin\theta$. This is similar to the field produced by an electric dipole in the $\hat{z}$ direction. In fact if we use the aperture field distribution to calculate the radiation fields, the equivalent surface current density will have no z component at all.

We therefore consider only the transverse component of $\overline{f}$ and write

$$\overline{f} = \iint dS\ \hat{e}_r \cos\frac{\theta'}{2} \frac{2E(\theta', \phi')}{\eta r'} e^{ikr' - i\overline{k} \cdot \overline{r}'} \tag{62}$$

The differential area dS on the paraboloidal surface is

$$dS = (r' \sin\theta'\, d\phi')(r' \sec(\frac{\theta'}{2})\, d\theta')$$

The exponential phase factor inside the integral is

$$-kr' + \overline{k} \cdot \overline{r}' = kr' \left[-1 + \sin\theta \sin\theta' \cos(\phi' - \phi) + \cos\theta \cos\theta'\right]$$

We now examine antenna gain along the z axis where $\theta = \pi$. The radiation vector in (62) becomes

$$\overline{f} = \int_0^{2\pi} d\phi' \int_0^{\theta_0} d\theta'\ \hat{e}_r r' \sin\theta' \frac{E(\theta', \phi')}{\eta} e^{i2kf} \tag{63}$$

Neglecting the cross-polarization effect and letting $\hat{e}_r \approx \hat{y}$, we obtain

$$\overline{E}(r,\pi,0) = \hat{y}\frac{i\omega\mu e^{ik(r+2f)}}{\pi\eta r}\int_0^{2\pi} d\phi' \int_0^{\theta_0} d\theta' f \tan\frac{\theta'}{2} E(\theta',\phi') \tag{64}$$

Suppose the gain of the feed is independent of ϕ'. For instance, a stub-supported dipole-feed [Silver, 1949] satisfies this requirement. Remembering that

$$\frac{1}{2\eta}|E(\theta',\phi')|^2 = \frac{P_t}{4\pi}G(\theta') \tag{65}$$

where P_t is the total radiated power by the feed, we find from (64) by integrating over ϕ'

$$\overline{E}(r,\pi,0) = \hat{y}\frac{i2\omega\mu f\, e^{ik(r+2f)}}{\eta r}\int_0^{\theta_0} d\theta' \tan\frac{\theta'}{2}\sqrt{\left[\frac{\eta P_t}{2\pi}G(\theta')\right]} \tag{66}$$

The antenna gain in the forward direction is

$$g = \frac{4\pi r^2}{P_t}\frac{1}{2\eta}|E(r,\pi,0)|^2 = 4k^2f^2\left[\int_0^{\theta_0} d\theta' \tan\frac{\theta'}{2}\sqrt{[G(\theta')]}\right]^2 \tag{67}$$

To pursue interpretations of (67), we define an aperture diameter D as [Fig. 5.7.12]

$$D = 2f\sec^2\frac{\theta_0}{2}\sin\theta_0 = 4f\tan\frac{\theta_0}{2} \tag{68}$$

We cast (67) in the form

$$g = \left(\frac{\pi D}{\lambda}\right)^2 \cot^2\frac{\theta_0}{2}\left[\int_0^{\theta_0} d\theta' \tan\frac{\theta_0}{2}[G(\theta')]^{1/2}\right]^2 \tag{69}$$

The factor $(\pi D/\lambda)^2$ is the gain for a uniformly illuminated circular aperture.

For a fixed D, the gain g increases as θ_0 increases because the available fraction of the total power from the feed is increased. But as θ_0 increases with a fixed D, the aperture efficiency with which the reflector concentrates the available power in the forward direction decreases because the illumination becomes less tapered from the center

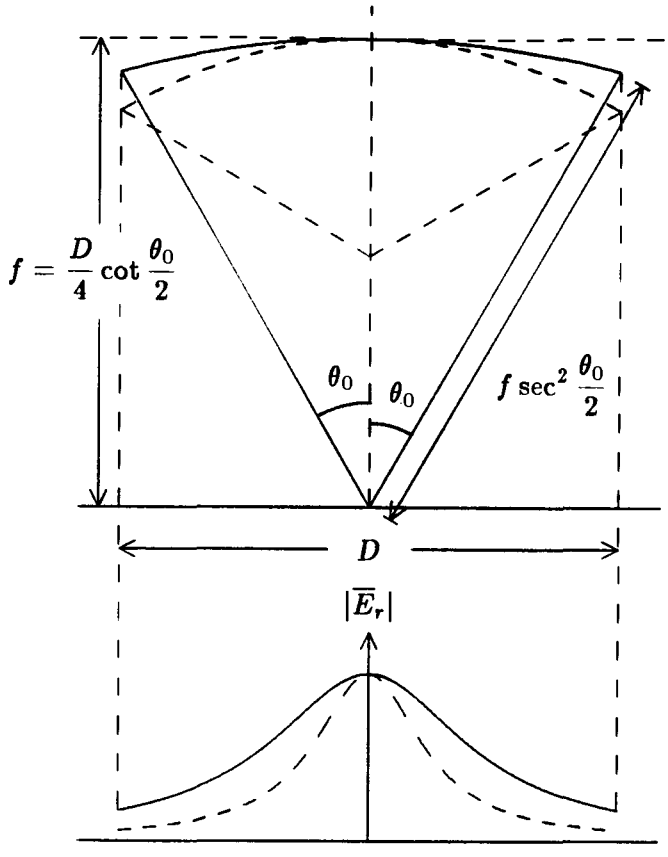


Figure 5.7.12 For a fixed D, gain increases as θ_0 increases.

[Fig. 5.7.12]. As a result, the gain g is decreased. The optimum angle θ_0 is obtained by setting $dg/d\theta_0 = 0$ which gives

$$\sqrt{[G(\theta_0)]} = \frac{1}{2}\csc^2\left(\frac{\theta_0}{2}\right)\int_0^{\theta_0} d\theta' \tan\frac{\theta'}{2}\sqrt{[G(\theta')]} \qquad (70)$$

The optimum angular aperture represents the proper compromise between spill-over of the feed energy and the aperture efficiency.

The gain can be modified by a number of factors: (i) phase-error effects due to deviation of the antenna-feed wavefronts from spherical ones, and defocusing effects due to displacement of the feed center from the focus. The maximum tolerable phase deviation is usually set at $\lambda/8$ over the aperture. (ii) The back lobe of the primary antenna

feed will interfere with the main lobe of the radiation pattern. This effect may be put into positive use by designing, according to the focal length f, a back lobe that interferes constructively with the main lobe along the axial direction. For a fixed-feed design it may be defocused to achieve constructive interference. (iii) The current at the edge of the reflector acts like a line source and radiates into the backward direction. Such currents can be reduced by making the edges irregular, by placing chokes or impedances on the edge, or by cutting properly spaced loops near the edge. (iv) The blockage of the feed in the path of the reflected ray significantly reduces the gain. This may be remedied by off-axis feeding arrangements but these create other electrical and hardware problems. (v) Cassegrain-fed paraboloidal reflector antennas with a hyperboloidal subreflector are sometimes used with the primary feeding source placed behind the main reflector. The arrangement permits the installation of complex primary feeds, reduces the temperature noise interference in radio astronomical receivers, and shortens the distance between the two reflectors because the subreflector will be placed before the focal point of the paraboloid. However, the reflector mismatch induced by spurious surface currents on the subreflector due to reflected waves from the main reflector may become serious.

5.8 Paraxial Limit

In the paraxial limit, solutions to the scalar wave equation

$$(\nabla^2 + k^2)U = 0 \tag{1}$$

can be written as

$$U = u(x, y, z)e^{ikz} \tag{2}$$

Under the paraxial approximation when we assume that $\partial^2 u/\partial z^2$ is negligible, the wave equation (1) becomes

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + izk\frac{\partial u}{\partial z} = 0 \tag{3}$$

We substitute into (3) the trial solution

$$u(x, y, z) = X\left(\frac{\sqrt{2}x}{w}\right) Y\left(\frac{\sqrt{2}y}{w}\right) e^{+i\left[p+\frac{k}{2q}(x^2+y^2)+\Phi(z)\right]} \tag{4}$$

where the parameters $w, p,$ and q are all functions of z.

a. Gaussian Beam

For the Gaussian beam solution, we let $X = Y = 1$ and $\Phi(z) = 0$. We find for

$$U(\rho, z) = e^{i\left[p(z) + \frac{k}{2q(z)}\rho^2\right]} \tag{5}$$

where $\rho^2 = x^2 + y^2$, Equation (3) yields

$$2k\left[-p' + \frac{i}{q}\right] + \frac{k^2\rho^2}{q^2}\left[q' - 1\right] = 0 \tag{6}$$

For (6) to be valid for all ρ and z, we obtain $p' = i/q$ and $q' = 1$. In the Gaussian beam expression, these conditions are made by choosing

$$q = z - i\frac{kw_o^2}{2} \tag{7}$$

$$p = \tan^{-1}\frac{2z}{kw_o^2} - i\ln\left(\frac{w_o}{w(z)}\right) \tag{8}$$

with

$$w(z) = w_o\sqrt{1 + \left(\frac{2z}{kw_o^2}\right)^2} \tag{9}$$

describing the beamwidth as a function of z.

The Gaussian beam solution thus takes the form, by virtue of (2), (5), (7) and (8)

$$u(\rho, z) = \frac{w_o}{w(z)}\exp\left[-i\tan^{-1}\frac{2z}{kw_o^2}\right]e^{ikz}\,e^{i\frac{2kz\rho^2}{4z^2 + (kw_o^2)^2}}\,e^{-\frac{(kw_o\rho)^2}{4z^2 + (kw_o^2)^2}} \tag{10}$$

The ρ-dependent exponentially decaying term $e^{-(kw_o\rho)^2/[4z^2 + (kw_o^2)^2]}$ describes the width of the Gaussian beam as a function of z. At $z = 0$, we define the width to be $\rho = w_o$ for which the term becomes e^{-1}. The ρ-dependent phase term $e^{i2kz\rho^2/[4z^2 + (kw_o^2)^2]}$ describes the phase front curvature. The phase shift at $\rho = \rho_o$ as compared to that at $\rho = 0$ is

$$kb = \frac{2kz\rho_o^2}{4z^2 + (kw_o^2)^2} \tag{11}$$

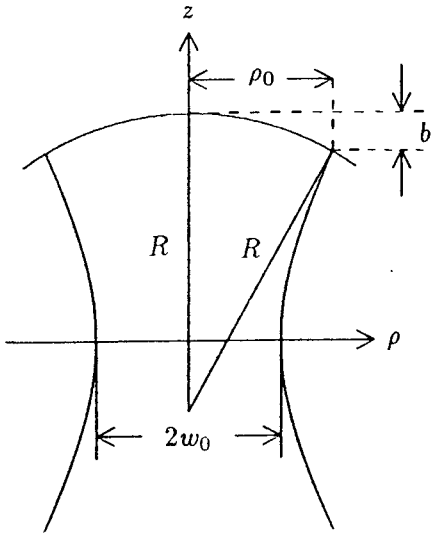


Figure 5.8.1

In the paraxial approximation, we have

$$\rho_o^2 + R^2 \approx (R+b)^2 \approx R^2 + 2Rb$$

From (11) we find

$$R = \frac{\rho_o^2}{2b} = z\left[1 + \frac{(kw_o^2)^2}{4z^2}\right] \tag{12}$$

At $z = 0$, the phase front curvature is infinite. Using (9) and (12) in (10), we obtain

$$U(\rho, z) = \frac{w_o}{w(z)} e^{i\left[kz + \frac{k\rho^2}{2R(z)} - \tan^{-1}\frac{2z}{kw_o^2}\right]} e^{-\left[\frac{\rho}{w(z)}\right]^2} \tag{13}$$

At $z = 0$, the beam waist is $\rho = w_o$ and the phase front is that of a plane wave. With increasing z, the beam waist increases, and the phase front curves with the curvature equal to R. As $z \to \infty$, the phase front curvature vanishes, while the arctangent terms gives $\pi/2$, and the phase front again approaches that of a plane wave.

The scalar wave equation in (1) can be taken to be the x component of a vector potential $\overline{A} = \hat{x}U(\rho, z)$. The Gaussian beam solution (13) is obtained from the paraxially-approximated equation (2) under the assumption that $|\partial u/\partial z| \ll |ku|$. Corresponding to the paraxial approximation, the magnetic field is

$$\overline{H} = \frac{1}{\mu}\nabla \times \overline{A} = ik\left[\hat{y}U + i\hat{z}\frac{i}{k}\frac{\partial U}{\partial y}\right]$$

and the electric field is

$$\overline{E} = iw\left[\hat{x}U + i\hat{z}\frac{1}{k}\frac{\partial U}{\partial x}\right]$$

and the total power in a Gaussian beam is

$$\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \frac{1}{2} Re\left\{\overline{E} \times \overline{H}^*\right\} \approx \hat{z}\frac{1}{2\eta}$$

where the amplitude of the Gaussian beam is assumed to be unity.

b. Gaussian-Hermite Beam Modes

Let the parameters p, q, and w in the trial solution (4) be given by (7)–(9). Substituting (4) into (3) and making use of (6), we find

$$\frac{X''}{X} - i\sqrt{2}\frac{kx}{w}\left(ww' - \frac{w^2}{q}\right)\frac{X'}{X} + \frac{Y''}{Y} - i\sqrt{2}\frac{kx}{w}\left(ww' - \frac{w^2}{q}\right) + kw^2\Phi'(z) = 0 \tag{14}$$

It follows from (7) and (9) that $ww' - w^2/q = -i2/k$. Separation of variables and setting

$$kw^2\Phi'(z) = 2(m+n) \tag{15}$$

lead to the differential equations for Hermite polynomials

$$X''\left(\frac{\sqrt{2}x}{w}\right) - 2\left(\frac{\sqrt{2}x}{w}\right)X' + 2mX = 0 \tag{16a}$$

$$Y''\left(\frac{\sqrt{2}y}{w}\right) - 2\left(\frac{\sqrt{2}y}{w}\right)Y' + 2nY = 0 \tag{16b}$$

Requiring $\Phi(z=0)=0$ and introducing (9), we obtain the solution for (15)

$$\Phi(z)=(m+n)\tan^{-1}\left(\frac{2z}{kw_o^2}\right) \tag{17}$$

The solution as expressed in (2) and (4) now takes the form, after use is made of (7), (8), and (13)

$$U=\frac{w_o}{w(z)}H_m\left(\frac{\sqrt{2}x}{w}\right)H_n\left(\frac{\sqrt{2}y}{w}\right)e^{-[\rho/w(z)]^2} \cdot e^{i\left[kz-(m+n+1)\tan^{-1}(2z/kw_o^2)+k\rho^2/2R(z)\right]} \tag{18}$$

where $H_m(\sqrt{2}x/w)$ and $H_n(\sqrt{2}y/w)$ are the Hermite polynonials of orders m and n.

Hermite polynomials $H_m(\xi)$ are solutions to the differential equation

$$\frac{d^2}{d\xi^2}H_m(\xi)-2\xi\frac{d}{d\xi}H_m(\xi)+2mH_m(\xi)=0 \tag{19}$$

The first five Hermite polynomials are

$$H_0(\xi)=1 \tag{20a}$$
$$H_1(\xi)=2\xi \tag{21b}$$
$$H_2(\xi)=4\xi^2-2 \tag{20c}$$
$$H_3(\xi)=8\xi^3-12\xi \tag{20d}$$
$$H_4(\xi)=16\xi^4-48\xi^2+12 \tag{20e}$$

The recurrence formula reads

$$H_{m+1}-2\xi H_m+2mH_{m-1}=0 \tag{21}$$

The Hermite polynomials can be defined by

$$H_m(\xi)=(-1)^m e^{\xi^2}\frac{d^m}{d\xi^m}e^{-\xi^2} \tag{22}$$

There also exist integral relations

$$i^n e^{-\xi^2/2}H_m(\xi)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}d\varsigma e^{i\xi\varsigma}e^{-\varsigma^2/2}H_m(\varsigma)$$

and the orthogonality conditions

$$\int_{-\infty}^{\infty} d\xi\, H_m(\xi) H_n(\rho) e^{-\xi^2} = 0$$

$$\int_{-\infty}^{\infty} d\xi\, H_m^2(\xi) e^{-\xi^2} = \sqrt{\pi} 2^n n!$$

From (18), we observe that all Gaussian-Hermite beam modes have the same beam width parameter $w(z)$, the same phase curvature $R(z)$, and the same q parameter in (7). It is to be noted that the Gaussian-Hermite beam modes form a complete system of orthogonal functions and thus paraxial wave fields can be expressed as a superposition of such modes. However the paraxial modes are only approximate solutions to the exact wave equation and the approximation becomes poorer as higher order modes become important.

c. Transmission of Gaussian Beams

The q parameter as defined in (7) has fundamental importance in the description of ray optics with Gaussian beams. In view of (9) for the definition of beamwidth w and (12) for the definition of curvature R, we can write

$$\frac{1}{q} = \frac{1}{z - ikw_0^2/2} = \frac{1}{R} + i\frac{2}{kw^2} \tag{23}$$

The q parameter at z becomes

$$q_1 = q + d \tag{24}$$

at $z + d$. A thin lens can be defined as a device which renders the Gaussian beam in (18) a phase shift $e^{-ik\rho^2/2f}$ after the beam passes the lens. We have from (18) the new curvature R_e after transmission to be

$$\frac{1}{R_e} = \frac{1}{R} - \frac{1}{f} \tag{25}$$

where f is the focal length of the lens. A mirror of radius R_0 acts like a lens with a focal length of $f = R_0/2$.

Both (24) and (25) can be expressed as a bilinear transformation with the form

$$q_1 = \frac{A_1 q + B_1}{C_1 q + D_1}$$

For (24)

$$\begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} \tag{26}$$

And for (25)

$$\frac{1}{q_l} = \frac{1}{q} - \frac{1}{f} \tag{27}$$

gives

$$q_l = \frac{q}{1 - q/f} = \frac{A_l q + B_l}{C_l q + D_l}$$

The corresponding $ABCD$ matrix is

$$\begin{bmatrix} A_l & B_l \\ C_l & D_l \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix} \tag{28}$$

For a beam with parameter q propagated over a distance d_1, transmitted through a lens of focal length f, and propagated over another distance d_2, we have the last parameter

$$\begin{aligned} q_2 = q_l + d_2 &= \frac{q_1 + d_2 - d_2 q_1/f}{1 - q_1/f} \\ &= \frac{(1 - d_2/f)q + d_1 + d_2 - d_1 d_2/f}{-q/f + 1 - d_1/f} = \frac{A_2 q + B_2}{C_2 q + D_2} \end{aligned}$$

The $A_2B_2C_2D_2$ matrix is easily shown to be a chain multiplication of the following three matrices:

$$\begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix} = \begin{bmatrix} 1 & d_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1/f & 1 \end{bmatrix} \begin{bmatrix} 1 & d_1 \\ 0 & 1 \end{bmatrix}$$

This chain rule of multiplication is characteristic of bilinear transformations and is extremely useful in the study of Gaussian beam transmission optics.

The $ABCD$ matrix is useful in the description of ray optics. Let a ray have beam position ρ_1 and slope ρ_1'. The ray passes through an optical system with rotational symmetry and acquires a new beam position ρ_2 and slope ρ_2' [Fig. 5.8.2]. We can write

$$\begin{bmatrix} \rho_2 \\ \rho_1' \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_1' \end{bmatrix} \tag{29}$$

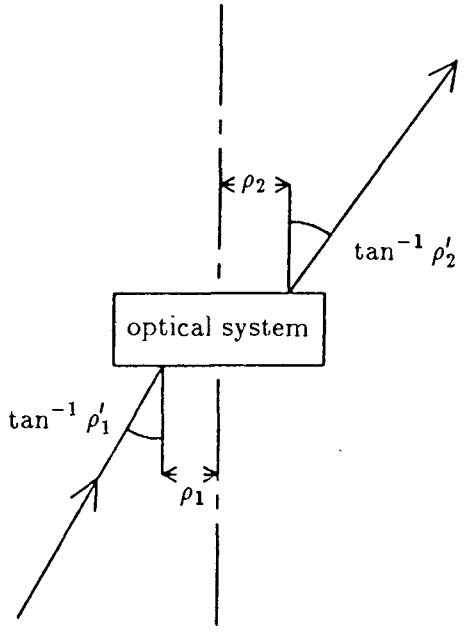


Figure 5.8.2 Transmission through an optical system.

For propagation in free space, the $ABCD$ matrix is identical to (26) and we find $\rho_2 = \rho_1 + d\rho_1'$ and $\rho_2' = \rho_1'$. For transmission through a lens, the $ABCD$ matrix is identical to (28) and we have $\rho_2 = \rho_1$ and $\rho_1' = \rho_1' - \rho_1/f$.

5.9 Quasi-Static Limits

The electromagnetic field vectors for time-harmonic fields are generally functions of the angular frequency ω. For example, the magnetic field $\overline{H}$ for a Hertzian dipole with dipole moment $Il = -i\omega p$ pointing in the $\hat{z}$ direction is

$$\overline{H} = \hat{\phi}\frac{-iIle^{ikr}}{4\pi R}\left[\frac{i}{kr} + \left(\frac{i}{kr}\right)^2\right]\sin\theta \tag{1}$$

where $i = \omega/c$. To explore the quasi-static limit, we expand the field vectors in power series of their values around $\omega = 0$. To facilitate the expansion, we separate the field dependence by introducing a new independent variable $\tau = \omega t$.

The field quantities $\overline{E}(\bar{r},\tau,\omega)$, $\overline{H}(\bar{r},\tau,\omega)$, $\overline{D}(\bar{r},\tau,\omega)$, $\overline{B}(\bar{r},\tau,\omega)$, $\overline{J}(\bar{r},\tau,\omega)$, and $\rho(\bar{r},\tau,\omega)$ can now be expanded in power series in ω. We write

$$\overline{E}(\bar{r},\tau,\omega) = \sum_{m=0}^{\infty} \frac{\omega^m}{m!} \left[\frac{\partial^m}{\partial\omega^m}\overline{E}(\bar{r},\tau,\omega)\right]_{\omega=0}$$

$$= \sum_{m=0}^{\infty} \overline{E}^{(m)}(\bar{r},\tau,\omega) \tag{2}$$

and similarly for other field quantities. From the Maxwell equations, we find

$$\sum_{m=0}^{\infty} \nabla\times\overline{E}^{(m)} = -\omega\frac{\partial}{\partial\tau}\sum_{m=0}^{\infty}\overline{B}^{(m)}$$

$$= -\sum_{m=1}^{\infty}\frac{\partial}{\partial\tau}\frac{\omega^m}{(m-1)!}\left[\frac{\partial^{m-1}}{\partial\omega^{m-1}}\overline{B}(\bar{r},\tau,\omega)\right]_{\omega=0}$$

$$= -\sum_{m=1}^{\infty}\omega\frac{\partial}{\partial\tau}\overline{B}^{(m-1)} \tag{3}$$

It follows that

$$\nabla\times\overline{E}^{(0)} = 0 \tag{4a}$$

$$\nabla\times\overline{E}^{(m)} = -\frac{\partial}{\partial t}\overline{B}^{(m-1)} \qquad m=1,2,\ldots \tag{4b}$$

Similarly,

$$\nabla\times\overline{H}^{(0)} = \overline{J}^{(0)} \tag{5a}$$

$$\nabla\times\overline{H}^{(m)} = \overline{J}^{(m)} + \frac{\partial}{\partial t}\overline{D}^{(m-1)} \qquad m=1,2,\ldots \tag{5b}$$

$$\nabla\cdot\overline{D}^{(0)} = \rho^{(0)} \tag{6a}$$

$$\nabla\cdot\overline{D}^{(m)} = \rho^{(m)} \qquad m=1,2,\ldots \tag{6b}$$

$$\nabla\cdot\overline{B}^{(0)} = 0 \tag{7a}$$

$$\nabla\cdot\overline{B}^{(m)} = 0 \qquad m=1,2,\ldots \tag{7b}$$

$$\nabla\cdot\overline{J}^{(0)} = 0 \tag{8a}$$

$$\nabla\cdot\overline{J}^{(m)} = -\frac{\partial}{\partial t}\rho^{(m-1)} \qquad m=1,2,\ldots \tag{8b}$$

Equations (4*a*) and (6*a*) are the governing equations for electrostatic fields. Equations (5*a*) and (7*a*) are the governing equations for magnetostatic fields.

It is important to notice that the mth-order fields are now expressed in terms of the $(m-1)$th-order fields. Thus, when the zeroth-order fields are obtained for given sources $\overline{J}^{(0)}$ and $\rho^{(0)}$, all higher order fields can be determined in succession. The electrostatic fields are governed by (4*a*), (6*a*), and (8*b*), with $m = 1$. Namely, the zeroth-order electrostatic fields as obtained from (4*a*) and (6*a*) are now used to find the first-order current $\overline{J}^{(1)}$ from (8*b*) and likewise the first-order magnetic field $\overline{H}^{(1)}$ from (4*b*). Equations (5*a*), (7*a*), and (4*b*) with $m = 1$ are seen to be the magnetostatic equations. Associated with each differential equation in (4)–(8), corresponding boundary conditions can be derived.

As an example, we see from (1) that the first-order magnetic field for the Hertzian dipole is

$$\overline{H}^{(0)} = \hat{\phi}\frac{Il}{4\pi r^2}\sin\theta$$

which is identical to the magnetic field for an infinitesimal current element pointing in the $\hat{z}$ direction (as obtained from the Biot-Savart law).

As another example, we consider the electromagnetic fields in a parallel-plate waveguide [Fig. 5.9.15]. A time-harmonic voltage source is applied at $z = 0$ with

$$V(t) = V_0\cos\omega t \tag{9}$$

The zeroth-order electric field is the electrostatic field

$$\overline{E}^{(0)} = -\hat{x}\frac{V_0}{d}\cos\omega t \tag{10a}$$

$$\overline{H}^{(0)} = 0 \tag{10b}$$

The surface charge density at $x = d$ is

$$\rho_s^{(0)} = \hat{n}\cdot\epsilon\overline{E}^{(0)} = \frac{\epsilon V_0}{d}\cos\omega t \tag{10c}$$

An opposite surface charge density is found on the plate at $x = 0$.

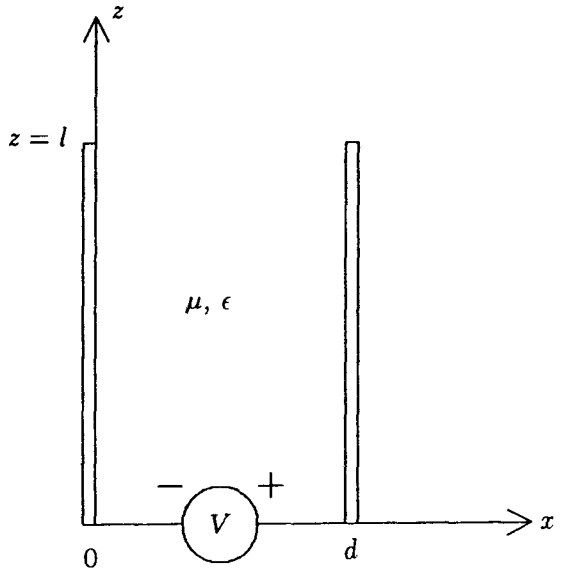


Figure 5.9.1 Parallel-plate waveguide.

The first-order surface current density follows from (8*b*),

$$\nabla \cdot \overline{J}_s^{(1)} = -\frac{\partial}{\partial t}\rho_s^{(0)} = \omega \frac{\epsilon V_0}{d} \sin \omega t \tag{11a}$$

Assuming that the plate dimension in the $\hat{y}$ direction is large so that the fields and currents are independent of y, the continuity equation (8*b*) becomes

$$\frac{\partial}{\partial z} J_s^{(1)} = \omega \frac{\epsilon V_0}{d} \sin \omega t$$

and we obtain

$$\overline{J}_s^{(1)} = \hat{z}\, \omega \frac{\epsilon V_0}{d} z \sin \omega t \tag{11b}$$

The first-order magnetic field $\overline{H}^{(1)}$ is found from (5*b*), which gives

$$\frac{\partial}{\partial z} H_y^{(1)} = -\omega \frac{\epsilon V_0}{d} \sin \omega t$$

Thus

$$\overline{H}^{(1)} = -\hat{y}\, \omega \frac{\epsilon V_0}{d} z \sin \omega t \tag{11c}$$

It is seen that (11*b*) and (11*c*) are related by the boundary condition at $x = d$

$$\overline{J}_s^{(1)} = \hat{n} \times \overline{H}^{(1)} = -\hat{x} \times \overline{H}^{(1)} \tag{11d}$$

Since $\overline{H}^{(0)} = 0$, we find from (74*b*) that

$$\overline{E}^{(1)} = 0 \tag{11e}$$

for the first-order electric field.

For the second-order field quantities, we find from (5*b*) that

$$\overline{H}^{(2)} = 0 \tag{12a}$$

which follows from (11*e*). As seen from (4*b*),

$$\nabla \times \overline{E}^{(2)} = \hat{y}\omega^2 \frac{\mu\epsilon V_0}{d} z \cos \omega t$$

Thus

$$\overline{E}^{(2)} = \hat{x} \frac{V_0}{d} \frac{k^2 z^2}{2} \cos \omega t \tag{12b}$$

where $k^2 = \omega^2 \mu\epsilon$.

Continuing the process, we find that the odd-order terms $\overline{E}^{(m)}$ are all zero and the even-order terms are

$$\overline{E}^{(2n)} = -\hat{x}(-1)^n \frac{V_0}{d} \frac{(kz)^{2n}}{(2n)!} \cos \omega t \qquad n = 0, 1, 2, \ldots \tag{13a}$$

Similarly, we find that the even-order terms for $\overline{H}^{(m)}$ are all zero and the odd-order terms are

$$\overline{H}^{(2n+1)} = -\hat{y}(-1)^n \frac{V_0}{\mu d} \frac{(kz)^{2n+1}}{(2n+1)!} \sin \omega t \qquad n = 0, 1, 2, \ldots \tag{13b}$$

Summing over the infinite number of terms, we obtain

$$\overline{E} = \sum_{n=0}^{\infty} \overline{E}^{(2n)} = -\hat{x} \frac{V_0}{d} \cos kz \cos \omega t \tag{14a}$$

$$\overline{H} = \sum_{n=0}^{\infty} \overline{H}^{(2n+1)} = -\hat{y} \frac{V_0}{\mu d} \sin kz \sin \omega t \tag{14b}$$

Equation (14) represents the exact solution for an open-circuited transmission line.

5.10 Quantization of Electromagnetic Waves

In microscopic physics, quantum electrodynamics has become a well-established discipline. In the study of interactions of electromagnetic waves with material media, semi-classical approaches are usually taken, as a full quantum theory is too complicated to carry out. We can either treat electromagnetic waves classically or treat material media classically. In this section we quantize electromagnetic waves with material media characterized by constitutive relations. The Heisenberg representation will be used in the process of quantization. First we give a brief review of this representation.

The state of a physical system is represented by a state vector, which can be viewed either as a column matrix called a *ket* and denoted by $|\psi>$ or as a row matrix called a *bra* and denoted by $<\psi|$. The bra $<\psi|$ is the complex conjugate and transpose of the ket $|\psi>$. Physical observables are represented by Hermitian operators. An operator can be viewed as a square hermitian matrix. Any measurement of a physical observable A yields a statistical expectation value for the observable. The average value of a series of measurements made on an ensemble of systems characterized by the same state vector $|\psi>$ is given by

$$< A > = < \psi | A | \psi >$$

This average value, or expectation value, is a real scalar number.

Each Hermitian operator possesses a set of eigenvectors with associated eigenvalues. The eigenvectors are eigenstates of the corresponding observable. The result of a single measurement for an observable A on a system described by the state vector $|\psi>$ yields an eigenvalue λ_n of the operator A and sends the system into the corresponding eigenstate $|\lambda_n>$. The probability of obtaining this eigenvalue λ_n and resulting in this eigenstate $|\lambda_n>$ is given by $|<\lambda_n|\psi>|^2$. As time progresses, the eigenstates of the operator evolve as the operator evolves with time, and the state of the system will no longer be an eigenstate of the operator. The evolution of the operator with time is determined by the equations of the motion. In classical electromagnetic theory the equations of motion for the electromagnetic field vectors are

Maxwell's equations. In quantum theory, the field vectors are treated as operators and are governed by Maxwell's equations.

a. Uncertainty Principle

Since operators do not necessarily commute, commutation relations for noncommuting operators must be postulated. The physical interpretation of the commutation relations leads to the uncertainty principle, which states, in essence, that any measurement made on a physical system, no matter how slight, perturbs the system. Thus reality is forever beyond reach; only statistical results show up. This perturbation, induced by a measurement on one observable, may or may not affect the true values of measurements on other observables. When two observables interfere with each other, they are noncommuting operators. When two observables do not interfere with each other, they are simultaneously measurable and the commutation relation is zero. Commutation relations for electromagnetic fields $\overline{D}$ and $\overline{B}$ are postulated as

$$\begin{aligned}[D_i(\bar{r},t), B_j(\bar{r}',t)] &= D_i(\bar{r},t)B_j(\bar{r}',t) - B_j(\bar{r}',t)D_i(\bar{r},t) \\ &= -i\hbar\,\epsilon_{ijk}\frac{\partial}{\partial x_k}\delta(\bar{r}-\bar{r}') \qquad (1)\end{aligned}$$

where $\hbar = 1.05 \times 10^{-34}$ joule-sec is Planck's constant divided by 2π. The factor $\hbar$ signifies that quantum effects are important whenever $\hbar$ is not numerically negligible. The classical limit is obtained when we let $\hbar \to 0$ and treat operators as classical variables.

In vacuum the commutation relations for other field components follow directly from the constitutive relations for vacuum:

$$[E_i(\bar{r},t), B_j(\bar{r}',t)] = -i\hbar\,\epsilon_{ijk}\frac{1}{\epsilon_o}\frac{\partial}{\partial x_k}\delta(\bar{r}-\bar{r}') \qquad (2)$$

$$[D_i(\bar{r},t), H_j(\bar{r}',t)] = -i\hbar\,\epsilon_{ijk}\frac{1}{\mu_o}\frac{\partial}{\partial x_k}\delta(\bar{r}-\bar{r}') \qquad (3)$$

$$[E_i(\bar{r},t), H_j(\bar{r}',t)] = -i\hbar c^2\,\epsilon_{ijk}\frac{\partial}{\partial x_k}\delta(\bar{r}-\bar{r}') \qquad (4)$$

The commutation relations state that the perpendicular components of the electric and magnetic fields interfere with each other, whereas parallel components are simultaneously measurable. For instance, to

measure an electric field, we may use a test charge and observe its motion along the electric field lines. But when a charge moves, it constitutes a current. The current produces a magnetic field perpendicular to the electric field. Thus the test charge used in the measurement of the electric fields interferes with a simultaneous measurement on magnetic fields.

Recall that the magnetic field $\overline{B}$ is expressible as the curl of a vector potential $\overline{A}$:

$$\overline{B} = \nabla \times \overline{A} \tag{5}$$

In terms of the vector potential $\overline{A}$, the commutation relation can be written as

$$[A_i(\overline{r},t), D_j(\overline{r}',t)] = -i\hbar\,\delta_{ij}\,\delta(\overline{r}-\overline{r}') \tag{6}$$

Note that, although all commutation regions are written for equal times, we can also deduce and postulate commutation relations for unequal times.

In our description of a quantized system, the operators evolve with time, as do their associated eigenvectors. The eigenvectors can be viewed as forming base vectors describing a system states vector that is not varying with time. This is known as the Heisenberg picture. The Heisenberg picture is different from the Schrödinger picture, in which the system state vectors are functions of time but the operators and their eigenstates are stationary. In the Schrödinger picture, the equation of motion of a state vector $|\psi(t)>$ is given by the Schrödinger equation. In the Heisenberg picture, the time evolution of an operator A representing a physical observable is governed by the Heisenberg equation of motion. Under the assumption that A is not explicitly dependent on time, the Heisenberg equation of motion for A is

$$i\hbar\frac{dA}{dt} = [A, \mathcal{H}] \tag{7}$$

where $\mathcal{H}$ is the Hamiltonian of the system, which corresponds to the total energy of the system. The Hamiltonian of an electromagnetic field in a source-free region is

$$\mathcal{H} = \int d^3\overline{r}\,\frac{1}{2}(\overline{E}\cdot\overline{D} + \overline{H}\cdot\overline{B}) \tag{8}$$

To determine the equation of motion of the $\overline{D}$ field, we use (7) and

find

$$i\hbar\frac{dD_i}{dt} = [D_i, \mathcal{H}]$$
$$= \frac{1}{2}\int d^3\bar{r}' \{[D_i(\bar{r}), E_j(\bar{r}')]D_j(\bar{r}') + E_j(\bar{r}')[D_i(\bar{r}), D_j(\bar{r}')]$$
$$+ [D_i(\bar{r}), H_j(\bar{r}')]B_j(\bar{r}') + H_j(\bar{r}')[D_i(\bar{r}), B_j(\bar{r}')]\}$$

Remember that all field observables are now operators. The integral can be evaluated with the use of commutation relations (1) and (3). The first two commutators are zero because electric field operators commute. We obtain

$$\frac{dD_i}{dt} = \epsilon_{ijk}\frac{\partial H_j}{\partial x_k} \tag{9}$$

which is Ampere's law in the absence of the source term J_i. If sources are present, the Hamiltonian in (8) must then include an interaction term $-\overline{J}\cdot\overline{A}$. In view of commutation relation (6), it is clear that this extra term gives rise to a source term J_i in (9). Following a similar procedure, we can derive Faraday's law from the Heisenberg equation of motion for $\overline{B}$.

b. Annihilation and Creation Operators

The eigenstates of the Hamiltonian are energy eigenstates because the Hamiltonian $\mathcal{H}$ is an energy operator. To facilitate discussion of the energy states of a quantized wave field, it is useful to transform the operator to $\overline{k}$ space:

$$\overline{D}(\bar{r}) = (2\pi)^{-3/2}\int d^3\overline{k}\ \overline{D}(\overline{k})e^{i\overline{k}\cdot\bar{r}} \tag{10a}$$

$$\overline{A}(\bar{r}) = (2\pi)^{-3/2}\int d^3\overline{k}\ \overline{A}(\overline{k})e^{i\overline{k}\cdot\bar{r}} \tag{10b}$$

The condition that $\overline{D}(\bar{r})$ and $\overline{A}(\bar{r})$ be real operators requires that

$$\overline{D}(-\overline{k}) = \overline{D}^+(\overline{k}) \tag{11a}$$

$$\overline{A}(-\overline{k}) = \overline{A}^+(\overline{k}) \tag{11b}$$

This reality condition can be satisfied by the following representations:

$$D_i(\overline{k}) = i\sqrt{\frac{\hbar k}{2\eta}}\,[a_i^+(\overline{k}) - a_i(-\overline{k})] \tag{12a}$$

$$A_i(\overline{k}) = \sqrt{\frac{\hbar\eta}{2k}}\,[a_i^+(\overline{k}) + a_i(-\overline{k})] \tag{12b}$$

where $\eta = \sqrt{\mu_o/\epsilon_o}$. As we shall demonstrate in subsequent developments, the operator $a_i(\bar{k})$ is an annihilation operator and the operator $a_i^+(\bar{k})$ is a creation operator. Upon operating on an energy eigenstate $a_i(\bar{k})$ annihilates a photon corresponding to wave vector $\bar{k}$ and polarization D_i in the state, whereas $a_i^+(\bar{k})$ creates such a photon.

The Hamiltonian $\mathcal{H}$ can be expressed as

$$\mathcal{H} = \frac{1}{2(2\pi)^3}\int d^3\bar{r}\, d^3\bar{k}\, d^3\bar{k}' \left\{\frac{1}{\epsilon_o}\overline{D}(\bar{k})\cdot\overline{D}(\bar{k}')e^{i(\bar{k}+\bar{k}')\cdot\bar{r}} - \frac{1}{\mu_o}[\bar{k}\times\overline{A}(\bar{k})]\cdot[\bar{k}'\times\overline{A}(\bar{k}')]e^{i(\bar{k}+\bar{k}')\cdot\bar{r}}\right\}$$

$$= \frac{1}{2}\int d^3\bar{k}\left\{\frac{1}{\epsilon_o}\overline{D}(\bar{k})\cdot\overline{D}^+(\bar{k}) + \frac{1}{\mu_o}[\bar{k}\times\overline{A}(\bar{k})]\cdot[\bar{k}\times\overline{A}^+(\bar{k})]\right\}$$

We focus our attention on a particular photon with a pre-specified $\bar{k}$ vector. The operator $\bar{a}$ and $\bar{a}^+$ are in the same direction as $\overline{D}$ and are perpendicular to $\bar{k}$. In terms of $\bar{a}(\bar{k})$ and $\bar{a}^+(\bar{k})$, the Hamiltonian becomes

$$\mathcal{H} = \frac{1}{2}\int d^3\bar{k}\ \hbar kc\,[a_i(\bar{k})a_i^+(\bar{k}) + a_i^+(\bar{k})a_i(\bar{k})] \tag{13}$$

and the commutation relation (6) becomes

$$i\frac{\hbar}{2}\left[\int d^3\bar{k}\,[a_i^+(\bar{k})e^{i\bar{k}\cdot\bar{r}} + a_i(k)e^{-i\bar{k}\cdot\bar{r}}],\right.$$

$$\left.\int d^3\bar{k}'\,[a_j^+(\bar{k}')e^{i\bar{k}'\cdot\bar{r}'} - a_j(\bar{k}')e^{-i\bar{k}'\cdot\bar{r}'}]\right] = -i\hbar\,\delta_{ij}\delta(\bar{r}-\bar{r}')$$

We deduce that

$$[a_i(\bar{k}), a_j^+(\bar{k}')] = \delta_{ij}\delta(\bar{k}-\bar{k}') \tag{14}$$

For the case $i = j$ and $\bar{k} = \bar{k}'$, we simply write

$$[a, a^+] = 1 \tag{15}$$

The Hamiltonian for the photon with a particular $\bar{k}$ vector becomes

$$\mathcal{H} = \frac{\hbar kc}{2}(a^+a + a^+a) = \hbar\omega(a^+a + \frac{1}{2}) \tag{16}$$

where we made the use of commutator (15) and the vacuum dispersion relation $\omega = kc$.

To obtain eigenvalues and eigenvectors for the energy operator $\mathcal{H}$, we write

$$\mathcal{H}\,|\mathcal{E}> = \mathcal{E}\,|\mathcal{E}> \tag{17}$$

where $|\mathcal{E}>$ denotes the eigenstate, and $\mathcal{E}$ the corresponding eigenvalue. We first show that the eigenvalue $\mathcal{E}$ is always non-negative. Scalar-multiplying (17) by the eigenbra $<\mathcal{E}|$ and using (16), we have

$$\hbar\omega <\mathcal{E}|\, a^+a + \frac{1}{2}\,|\mathcal{E}> = \mathcal{E}\ <\mathcal{E}|\,\mathcal{E}>$$

The scalar $<\mathcal{E}\ |\ \mathcal{E}>$ is always non-negative because it is the product of a column matrix $|\mathcal{E}>$ and its complex conjugate and transpose $<\mathcal{E}|$. The term $<\mathcal{E}\,|a^+a|\,\mathcal{E}>$ is also non-negative for the same reason. Note that $<\mathcal{E}\ |\ a^+$ is the complex conjugate and transpose of $a\ |\ \mathcal{E}>$. Consequently, the eigenvalue $\mathcal{E}$ must be non-negative.

We next show that, if $|\mathcal{E}>$ is an eigenstate of $\mathcal{H}$, so are $a\,|\,\mathcal{E}>$ and $a^+\,|\,\mathcal{E}>$. Consider $\mathcal{H}(a^+\,|\,\mathcal{E}>)$. Using commutation relation (15), we find

$$[\mathcal{H}, a^+] = \hbar\omega[a^+a, a^+] = \hbar\omega a^+$$

Thus

$$\mathcal{H}a^+\,|\,\mathcal{E}> = a^+\mathcal{H}\,|\,\mathcal{E}> + \hbar\omega a^+\,|\,\mathcal{E}> = (\mathcal{E}+\hbar\omega)a^+\,|\,\mathcal{E}> \tag{18}$$

It is seen that $a^+\,|\,\mathcal{E}>$ is an eigenstate of $\mathcal{H}$ with eigenvalue $(\mathcal{E}+\hbar\omega)$. Whenever a^+ is applied to an eigenstate of $\mathcal{H}$ with energy $\mathcal{E}$, the state changes into another eigenstate with energy $\mathcal{E}+\hbar\omega$. Since the net effect of this operation is to create one more photon with energy $\hbar\omega$, the operator a^+ is called a creation operator.

Following similar reasoning, we can show that

$$\mathcal{H}a\,|\,\mathcal{E}> = (\mathcal{E}-\hbar\omega)a\,|\,\mathcal{E}> \tag{19}$$

When a is applied to an eigenstate $|\mathcal{E}>$, the result is another eigenstate, $|\mathcal{E}-\hbar\omega>$, with one photon annihilated. The operator a is thus called an annihilation operator.

We have proved that all energy states of $\mathcal{H}$ possess non-negative energy eigenvalues. Suppose that we apply the annihilation operator a to the state $|\mathcal{E}>$ n times and reach the ground state $|\mathcal{E}_0>$. Further operation of a on $|\mathcal{E}_0>$ will then yield a zero:

$$a\,|\,\mathcal{E}_0> = 0$$

We find the energy of the ground state to be

$$\mathcal{E}_0 = \frac{<\mathcal{E}_0|\mathcal{H}|\mathcal{E}_0>}{<\mathcal{E}_0|\mathcal{E}_0>} = \frac{1}{2}\hbar\omega \tag{20}$$

The separation between energy levels is $\hbar\omega$. When operated on by a, the transition is downward; when operated on by a^+, the transition is upward. The whole energy spectrum can be built up by successively applying a^+ to the ground state, which we denote by ket $|0>$ with 0 indicating no photon in the state. The state of n photons, $|n>$, is then created by operating a^+ on $|0>$ n times. The energy eigenvalue associated with $|n>$ is clearly $(n+1/2)\hbar\omega$. The energy eigenstate $|n>$ can be represented by a column matrix with all elements equal to zero except the $(n+1)$ th one, which is equal to unity. For instance,

$$|0> = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

$$|1> = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

The Hamiltonian has been diagonalized. In matrix representation, we can write

$$\mathcal{H} = \hbar\omega \begin{bmatrix} 1/2 & 0 & 0 & \cdots \\ 0 & 3/2 & 0 & \cdots \\ 0 & 0 & 5/2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{21}$$

which is a diagonal matrix. Since

$$\mathcal{H}\,|n> = \hbar\omega\left(a^+a + \frac{1}{2}\right)|n> = \hbar\omega\left(n + \frac{1}{2}\right)|n>$$

we call $N = a^+a$ the number operator,

$$N\,|\,n> = n\,|\,n> \tag{22}$$

The eigenstates of $\mathcal{H}$ are also eigenstates of the number operator. The eigenvalue associated with a particular state of N is equal to the number of photons in that state.

We now find an explicit representation of a and a^+ in terms of matrices. We write

$$a \mid n > = C_n \mid n-1 > \tag{23}$$

The coefficient C_n is determined from normalization. We require that the scalar product of the bra and the ket of an eigenstate be unity. We have

$$n = < n \left| a^+ a \right| n > = |C_n|^2$$

Thus,

$$a \mid n > = \sqrt{n} \mid n-1 > \tag{24}$$

The matrix elements of a can be obtained from (24) by noting that

$$< n-1 \left| a \right| n > = \sqrt{n} \tag{25}$$

which is the element in the $(n-1)$ th row and the n th column. All other elements are zero:

$$a = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & \sqrt{2} & 0 & \dots \\ 0 & 0 & 0 & \sqrt{3} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{26}$$

Following similar reasoning, we let $a^+ \mid n > = C_n' \mid n+1 >$ and find

$$\left|C_n'\right|^2 = < n \left| a a^+ \right| n > = < n \left| a^+ a + 1 \right| n > = n+1$$

Thus,

$$a^+ \mid n > = \sqrt{n+1} \mid n+1 > \tag{27}$$

Forming the product

$$< n+1 \left| a^+ \right| n > = \sqrt{n+1} \tag{28}$$

we see that the matrix representation of a^+ takes the following form:

$$a^+ = \begin{bmatrix} 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & \dots \\ 0 & \sqrt{2} & 0 & 0 & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{29}$$

Obviously, the matrix representation for a^+ is the transpose of that for a. The time evolution of a^+ follows the Heisenberg equation of motion as $\overline{D}$ and $\overline{A}$ do. From matrix multiplication we see that operating a on the state $|n>$ will move the unit element in the nth position to the $(n+1)$th position and form the state $|n-1>$ multiplied by $\sqrt{n}$. Similarly, operating a^+ on $|n-1>$ results in the state $|n>$ multiplied by $\sqrt{n}$. Operating on $|n>$ by a^+a will result in the same state and give the photon number n.

We have discussed energy eigenstates and their associated eigenvalues for the Hamiltonian operator $\mathcal{H}$, and we have seen what results the annihilation operator a and creation operator a^+ have when operating on the eigenstates. It is natural to ask, "What are the eigenstates and eigenvalues for the annihilation and the creation operators ?" First we shall prove that a^+ has no nonzero eigenstates. Denote the eigenstates of a^+ as $|e>$ and let the eigenvalues be λ. In view of (29), we have

$$\lambda \begin{bmatrix} e_0 \\ e_1 \\ e_2 \\ e_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{3} & 0 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} e_0 \\ e_1 \\ e_2 \\ e_3 \\ \vdots \end{bmatrix}$$

We see that if $\lambda = 0$, then $e_1 = e_2 = \cdots = 0$ and $|e>=|0>$. If $\lambda \neq 0$, then $e_0 = 0, e_1 = (1/\lambda)e_0 = 0, \ldots, e_n = (1/\lambda)\sqrt{n}e_{n-1} = 0, \ldots$, and the eigenstate is identically zero.

The same procedure can be used to find eigenstates for a. We write, in view of (26),

$$\lambda \begin{bmatrix} e_0 \\ e_1 \\ e_2 \\ e_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{2} & 0 & \cdots \\ 0 & 0 & 0 & \sqrt{3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} e_0 \\ e_1 \\ e_2 \\ e_3 \\ \vdots \end{bmatrix} \tag{30}$$

We see that $e_1 = \lambda e_0, e_2 = (\lambda/\sqrt{2})e_1, \ldots, e_n = (\lambda/\sqrt{n})e_{n-1}, \ldots$ In terms of the energy eigenstates, we obtain

$$|e> = e_0 \left[|0> + \lambda |1> + \frac{\lambda^2}{\sqrt{2!}} |2> + \cdots + \frac{\lambda^n}{\sqrt{n!}} |n> + \cdots \right] \tag{31}$$

Imposing the normalization condition $< e \mid e > = 1$ yields

$$|e_0|^2 \left[1 + \lambda^2 + \frac{\lambda^4}{2!} + \cdots + \frac{\lambda^{2n}}{n!} + \cdots\right] = 1 \tag{32}$$

Thus $e_0 = \exp(-\lambda^2/2)$ and (31) becomes

$$|e> = e^{-\lambda^2/2} \sum_{n=0}^{\infty} \frac{\lambda^n}{\sqrt{n!}} \mid n> \tag{33}$$

We note that the expectation value for the photon number operator a^+a is determined from

$$\overline{n} = < e \left| a^+ a \right| e > = e^{-\lambda^2} \sum_{n=0}^{\infty} n \frac{\lambda^{2n}}{n!} = \lambda^2 \tag{34}$$

Consequently, the eigenvalue λ is equal to the square root of the photon number expectation value. Equation (33) becomes

$$|e> = e^{-\overline{n}/2} \sum_{n=0}^{\infty} \frac{\overline{n}^{n/2}}{\sqrt{n!}} \mid n> \tag{35}$$

This represents the eigenstate of the annihilation operator a. It is also called the coherent state. The probability of finding the average photon number of $\overline{n}$ is

$$|< n \mid e >|^2 = \frac{\overline{n}^{\overline{n}} e^{-\overline{n}}}{n!} \tag{36}$$

This is the Poisson distribution. It is seen that in the energy state representation the precise photon number is given, whereas in the coherent state representation the photon number obeys the Poisson probability distribution.

c. *Wave Quantization in Bianisotropic Media*

With the use of the annihilation and creation operators a and a^+, we have diagonalized the Hamiltonian for an electromagnetic field in vacuum. The quantized fields have been discussed in terms of the energy states. We now generalize the procedure to carry out wave quantization in bianisotropic media. The commutation relations for $\overline{D}$ and

$\overline{B}$ are the same as postulated in (1), and those for other field operators are derived from (1) by using the constitutive relations. Instead of (12), the annihilation and creation operators are introduced by

$$D_j(\overline{k}) = \frac{i}{\alpha_j}\sqrt{\frac{\hbar}{2}}\,[a_j^+(\overline{k}) - a_j(-\overline{k})] \tag{37a}$$

$$A_j(\overline{k}) = \alpha_j\sqrt{\frac{\hbar}{2}}\,[a_j^+(\overline{k}) + a_j(-\overline{k})] \tag{37b}$$

where α_j is a constant to be determined. In view of the reality condition (11), we must have $\alpha_j(-\overline{k}) = \alpha_j^*(\overline{k})$. Substituting (27) into (6), we find that the commutation relations for a and a^+ are identical to (14) and (15). Using a and a^+, we shall diagonalize the Hamiltonian by properly choosing α_j.

As an example of bianisotropic media we consider a uniaxial medium moving along the direction of its optic axis. We let the optic axis be along the $\hat{z}$ direction [see Chapter VII]. In the kDB system, $D_3 = B_3 = 0$ and the constitutive relation is

$$\overline{\overline{\kappa}}_k = \begin{bmatrix} \kappa & 0 & 0 \\ 0 & \kappa\cos^2\theta + \kappa_z\sin^2\theta & (\kappa-\kappa_z)\sin\theta\cos\theta \\ 0 & (\kappa-\kappa_z)\sin\theta\cos\theta & \kappa\sin^2\theta + k_z\cos^2\theta \end{bmatrix}$$

$$\overline{\overline{\nu}}_k = \begin{bmatrix} \nu & 0 & 0 \\ 0 & \nu\cos^2\theta + \nu_z\sin^2\theta & (\nu-\nu_z)\sin\theta\cos\theta \\ 0 & (\nu-\nu_z)\sin\theta\cos\theta & \nu\sin^2\theta + \nu_z\cos^2\theta \end{bmatrix}$$

$$\overline{\overline{\chi}}_k = \overline{\overline{\nu}}_k^+ = \begin{bmatrix} 0 & \chi\cos\theta & \chi\sin\theta \\ -\chi\cos\theta & 0 & 0 \\ -\chi\sin\theta & 0 & 0 \end{bmatrix}$$

The Hamiltonian becomes

$$\begin{aligned} \mathcal{H} =& \frac{1}{2}\int d^3\overline{k}\,\left[\overline{D}^+(\overline{k})\cdot\overline{\overline{E}}(\overline{k}) + \overline{B}^+(\overline{k})\cdot\overline{H}(\overline{k})\right] \\ =& \frac{1}{2}\int d^3\overline{k}\,\Big\{\Big[\kappa D_1^+(\overline{k})D_1(\overline{k}) + (\kappa\cos^2\theta + \kappa_z\sin^2\theta)D_2^+(\overline{k})D_2(\overline{k}) \\ &+ k^2\nu A_2^+(\overline{k})A_2(\overline{k}) + k^2(\nu\cos^2\theta + \nu_z\sin^2\theta)A_1^+(\overline{k})A_1(\overline{k})\Big] \\ &- ik\chi\cos\theta\Big[D_1^+(\overline{k})\overline{A}_1(\overline{k}) - A_1^+(\overline{k})D_1(\overline{k}) \\ &\qquad + D_2^+(\overline{k})A_2(\overline{k}) - A_2^+(\overline{k})D_2(\overline{k})\Big]\Big\} \end{aligned} \tag{38}$$

where θ is the angle between the $\bar{k}$ vector and the z axis. To express the Hamiltonian in terms of the annihilation and the creation operators, we note that

$$\int d^3\bar{k}\, k\{D_j^+(\bar{k})A_j(\bar{k}) - A_j^+(\bar{k})D_j(\bar{k})\} = i\hbar \int d^3k\, k\{a_j(\bar{k})a_j^+(\bar{k}) + a_j^+(\bar{k})a_j(\bar{k})\}$$

Introducing (37) in (38), we obtain

$$\begin{aligned}\mathcal{H} = \frac{\hbar}{4}\int d^3\bar{k}\Bigg\{&\left(\frac{\kappa}{\alpha_1^2} + k^2\alpha_1^2(\nu\cos^2\theta + \nu_z\sin^2\theta) + 2k\chi\cos\theta\right)\\ &\cdot [a_1(\bar{k})a_1^+(\bar{k}) + a_1^+(\bar{k})a_1(\bar{k})] + \left(\frac{1}{\alpha_2^2}(\kappa\cos^2\theta + \kappa_z\sin^2\theta)\right.\\ &\left.+ \kappa^2\alpha_2^2 + 2k\chi\cos\theta\right)[a_2(\bar{k})a_2^+(\bar{k}) + a_2^+(\bar{k})a_2(\bar{k})]\\ &+ \left(k^2\alpha_1^2(\nu\cos^2\theta + \nu_z\sin^2\theta) - \frac{\kappa}{\alpha_1^2}\right)[a_1(\bar{k})a_1(-\bar{k})\\ &+ a_1^+(-\bar{k})a_1^+(\bar{k})] + \left(k^2\nu\alpha_2^2 - \frac{1}{\alpha_2^2}(\kappa\cos^2\theta + \kappa_z\sin^2\theta)\right)\\ &\cdot [a_2(\bar{k})a_2(-\bar{k}) + a_2^+(-\bar{k})a_2^+(\bar{k})]\Bigg\}\end{aligned} \tag{39}$$

We see that the last two terms can be made to vanish by choosing

$$\alpha_1^4 = \frac{\kappa}{k^2(\nu\cos^2\theta + \nu_z\sin^2\theta)} \tag{40a}$$

$$\alpha_2^4 = \frac{(\kappa\cos^2\theta + \kappa_z\sin^2\theta)}{k^2\nu} \tag{40b}$$

The Hamiltonian $\mathcal{H}$ is seen to be diagonalized. It can be written as the sum of two Hamiltonians, each corresponding to a characteristic wave in the moving medium:

$$\mathcal{H} = \mathcal{H}_m + \mathcal{H}_e \tag{41a}$$

$$\mathcal{H}_m = \frac{\hbar}{2}\int d^3\bar{k}\left(\frac{\kappa}{\alpha_1^2} + k\chi\cos\theta\right)[a_1(\bar{k})a_1^+(\bar{k}) + a_1^+(\bar{k})a_1(\bar{k})] \tag{41b}$$

$$\mathcal{H}_e = \frac{\hbar}{2}\int d^3\bar{k}\left[\frac{(\kappa\cos^2\theta + \kappa_z\sin^2\theta)}{\alpha_2^2} + k\chi\cos\theta\right] \cdot [a_2(\bar{k})a_2^+(\bar{k}) + a_2^+(\bar{k})a_2(\bar{k})] \tag{41c}$$

In the case of a stationary uniaxial dielectric medium, $\chi = 0$. We see that the photons associated with $\mathcal{H}_m$ are the ordinary photons and those associated with $\mathcal{H}_e$ are the extraordinary photons. The Hamiltonians in (41) can be expressed in terms of the number operators. When the Hamiltonian is operated on an energy state, the result is the total photon energy in the state. The photon energies of the two types of photons corresponding to $\mathcal{H}_m$ and $\mathcal{H}_e$ are as follows:

$$E_m = \hbar\left(\frac{\kappa}{\alpha_1^2} + \kappa\chi\cos\theta\right) \tag{42a}$$

$$E_e = \hbar\left[\frac{(\kappa\cos^2\theta + \kappa_z\sin^2\theta)}{\alpha_2^2} + k\chi\cos\theta\right] \tag{42b}$$

The photon energies can be negative when the medium velocity is sufficiently high. They correspond to classical slow waves in moving media which can also possess negative energy, and this is a purely kinematic effect when transforming from the moving to the stationary frame.

In general, the energy of a photon with a specified wave vector $\overline{k}$ is equal to $\hbar$ times the angular frequency. Classically, the angular frequency is related to the $\overline{k}$ vector by dispersion relations. The derivation of the dispersion relations is facilitated by the use of the kDB system. Since each characteristic wave has a particular dispersion relation, we expect that for each characteristic wave there will be a corresponding photon after quantization is carried out.

PROBLEMS

Problem P5.1

Let the tangential $\overline{E}$ field and tangential $\overline{H}$ field on an enclosed surface S be related by an impedance matrix $\overline{\overline{Z}}(\overline{r})$ such that

$$\hat{n} \times \overline{E} = \overline{\overline{Z}}(\overline{r}) \cdot (\hat{n} \times \overline{H})$$

where $\hat{n}$ is inward normal to the surface S bounding the region of interest V

$$Z_{ij}(\overline{r}) = \hat{s}_i Z_{ij}(\overline{r})\hat{s}_j$$

and $\hat{s}_i$ and $\hat{s}_j$ with $i, j = 1, 2$ are unit vectors tangential to the surface S such that $\hat{s}_1, \hat{s}_2$, and $\hat{n}$ form an orthogonal coordinate system. Show that the uniqueness theorem holds if $Z_{11} = Z_{22}^*$, $Z_{12} = -Z_{12}^*$, and $Z_{21} = -Z_{21}^*$.

Problem P5.2

Calculate the diffraction pattern produced by a plane wave normally incident on a rectangular slit.

Problem P5.3

Using the equivalence principle, calculate the field radiated from the open end of a coaxial line. Assume that the field inside the waveguide has only TEM modes.

Problem P5.4

Consider a plane wave polarized in the $\hat{y}$ direction and incident on a sphere of radius a. The radius of the sphere is much larger than a wavelength. Assume that the following approximate values hold for the scattered fields $\overline{E}_s$ and $\overline{H}_s$ on the surface of the sphere: (i) in the shadow region $\overline{E}_s \approx -\overline{E}_i$ and $\overline{H}_s \approx -\overline{H}_i$, and (ii) in the illuminated region $\hat{r} \times \overline{E}_s \approx -\hat{r} \times \overline{E}_i$ and $\hat{r} \times \overline{H}_s \approx \hat{r} \times \overline{H}_i$, where $\overline{E}_i$ and $\overline{H}_i$ are the fields of the incident wave. This assumed source distribution is a result of the physical optics approximation. Solve for the scattered radiation fields and calculate the echo area, defined by

$$A_e = \lim_{r \to \infty} \left(4\pi r^2 \frac{P_s}{P_i}\right)$$

where P_s is the back scattered power density, and P_i is the incident power density of the plane wave.

Problem P5.5

A slot antenna consists of a slot opening in a metallic plane and has an infinitesimally small width w and length $2l$ in the $\hat{z}$ direction. Let the voltage across the slot at $z = 0$ be V and the electric field distribution in the slot region be $(V/w) \sin k(l - |z|)$. Let $l = \lambda/4$, show that the complementary structure is a wire antenna of length $\lambda/2$. Find the input impedance of the slot antenna.

Problem P5.6

Calculate the radiated field from an open parallel-plate waveguide. Assume that the field inside is the TE_1 mode.

Problem P5.7

Consider a plane wave incident on a conducting plate of length a in the $\hat{x}$ direction and width b in the $\hat{y}$ direction. Assume that the plate size is very large and that the field is negligible behind the plate. Solve the scattered field by two approximate source distributions.

(a) The tangential electric field at the conducting plate is zero. Under the physical optics approximation, the tangential magnetic field at the conducting plate is twice that of the incident field, so that

$$\overline{J}_s = 2\hat{n} \times \overline{H}^i = \hat{x}\frac{2E_0}{\eta}$$

Show that the scattered electric field is

$$\overline{E}_s = (\hat{\phi}\sin\phi - \hat{\theta}\cos\theta\cos\phi)\frac{2ikE_0e^{ikr}}{\pi r}\frac{\sin(k_xa/2)}{k_x}\frac{\sin(k_yb/2)}{k_y}$$

(b) The induction theorem makes use of the source distribution for the scattered field, given by $\overline{J}_s = -\hat{n} \times (\overline{H} - \overline{H}_i) = -\hat{n} \times \overline{H}_i$ and $\overline{M}_s = -\hat{n} \times (\overline{E} - \overline{E}_i) = \hat{n} \times \overline{E}_i = \hat{y}E_0$, where $\overline{E}$ and $\overline{H}$ represent the total field at the scatterer. These are impressed currents that radiate in the presence of the conductor. By the image theorem, the source responsible for the scattered field in the absence of the scatterer is $2M_s$, since the total field $\overline{E}$ at the conducting plate is zero and $\overline{J}_s$ does not radiate in front of perfect conductors. Show that the radiated electric field vector is

$$\overline{E}_s = (\hat{\phi}\cos\theta\sin\phi - \hat{\theta}\cos\phi)\frac{2ikE_0e^{ikr}}{\pi r}\frac{\sin(k_xa/2)}{k_x}\frac{\sin(k_yb/2)}{k_y}$$

(c) Compare the result obtained with the two approaches. Show that the results are identical in the backscattered direction. For a plane wave normally incident on a large plate, the backscattering cross-section, or radar cross-section, or echo area A_e is

$$A_e \equiv \lim_{r\to\infty}\left(9\pi r^2\frac{S^s}{S^i}\right) = \frac{k^2(\text{Area})^2}{\pi}$$

where $S^i = |E_o|^2/2\eta$ and S^s is the scattered power density.

Problem P5.8

By the image theorem, a vertical monopole antenna on a conducting plane is equivalent to a dipole with the conductor removed. In radio broadcasting stations, the Earth is used as the conducting plane. Calculate the power, the gain, and the radiation pattern for a monopole on a conducting plane.

Problem P5.9

Find multiple images for a dipole source in a parallel-plate wave guide and in a rectangular metallic waveguide.

Problem P5.10

Consider the problem of diffraction by a slit of width $2l$ in the $\hat{x}$ direction and infinite length in the $\hat{y}$ direction. A plane wave polarized in the $\hat{y}$ direction is normally incident upon the slit from $z \leq 0$. At the slit aperture, assume the field to be

$$\overline{E}(x, z = 0) = \hat{y} E_0 \, U(l - |x|)$$

where $U(l - |x|)$ is a unit step function. The field at any distance z can be viewed as a superposition of plane waves with the wave vector component k_x spanning $-\infty$ to ∞,

$$\overline{E}(x, z) = \hat{y} \int_{-\infty}^{\infty} dk_x \, \mathcal{E}(k_x) e^{ik_x x + ik_z z}$$

where $k_z = (k^2 - k_x^2)^{1/2}$. Show that at the aperture where $z = 0$

$$\mathcal{E}(k_x) = \frac{E_0 l}{\pi} \frac{\sin k_x l}{k_x l}$$

The major contribution to the integral comes from the interval $[-\pi/l, \pi/l]$ on the k_x axis, where the peak of $\mathcal{E}(k_x)$ is located. Assume $k \gg \pi/l$. Then, $k_x \ll k$ in the region where the integrand is significantly large. Approximate

$$k_z = (k^2 - k_x^2)^{1/2} \approx k - \frac{k_x^2}{2k} + \cdots$$

Show that the far field as evaluated by the stationary-phase method yields

$$\overline{E}(x,z) = \hat{y}e^{ikz}e^{i(kx^2/2z)}\int_{-\infty}^{\infty} dk_x\, \mathcal{E}(k_x)e^{-i(z/2k)[k_x-(kx/z)]^2}$$

$$\approx \hat{y}e^{ikz}e^{i(kx^2/2z)}\sqrt{2\pi k/iz}\;\mathcal{E}(kx/z)$$

Thus the far-field pattern is proportional to the Fourier transformation of the aperture field pattern. Show that for uniform illumination at the aperture,

$$\overline{E}(x,z) = \hat{y}E_0 l\sqrt{2k/i\pi z}\,\frac{\sin(klx/z)}{klx/z}e^{i(kx^2/2z)}e^{ikz}$$

For the screen located at a distance z, the phase at x is $kz + k(z^2 + x^2)^{1/2} \approx kz + kx^2/2z$ which explains the term $\exp(ikx^2/2z)$. The first zero is seen to occur at $klx_0/z = \pi$ which amounts to approximating the aperture field by two equivalent source centers a distance l apart such that they interfere destructively in the far field at an angle of $\theta_0 \approx x_0/z$. Sketch the radiation field at a constant distance $z = z_0$.

Problem P5.11

A Fabry-Perot resonator is composed of two parallel reflectors within which a standing wave is formed. The field inside is a TEM wave standing between the two plates. Show that the resonant wavenumbers are $k_z d = m\pi$. Strictly speaking, because of the finite transverse dimension, a TEM wave is diffracted. Considering diffraction in the cavity, we assume that the transverse field distribution takes a Gaussian form. At $z = 0$, $\overline{E} = \hat{x}E_0\, e^{-y^2/\omega_0^2}$. At $z > 0$, the electric field can be written as a superposition of plane waves with $\overline{k} = \hat{z}k_z + \hat{y}k_y$ and $k_y \ll k_z$:

$$\overline{E} = \hat{x}\int_{-\infty}^{\infty} dk_y\, E_g\, e^{ik_z z + ik_y y}$$

where

$$k_z = \sqrt{k^2 - k_y^2} \approx k\left(1 - \frac{1}{2}\frac{k_y^2}{k^2}\right)$$

The amplitude E_g is determined by the field at $z = 0$. Determine E_g by using an inverse Fourier transformation. Calculate and show that

$$\overline{E} \approx \hat{x}\frac{E_0}{\sqrt{1 + iz/z_F}}\, e^{ikz}\, e^{(iz/z_F - 1)y^2/w^2}$$

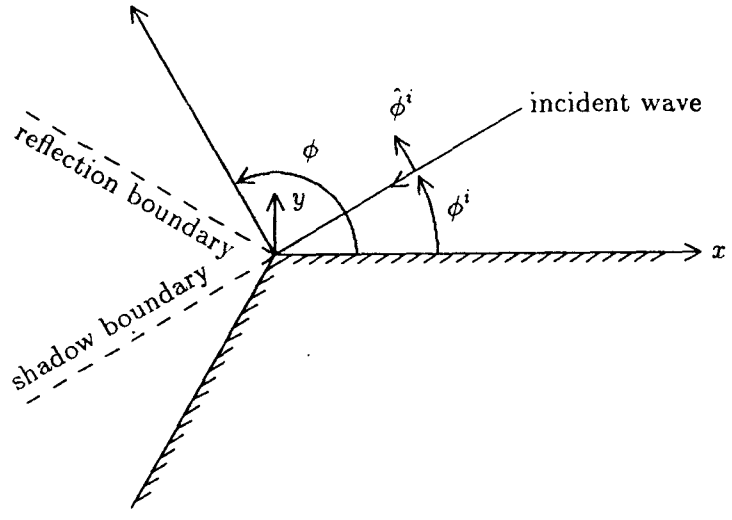


Figure P5.12a

where $z_F = kw_0^2/2$ and $w(z) = w_0\sqrt{1 + z^2/z_F^2}$. Prove that, for a given ω_0, the locus of w versus z is a hyperbola. The phase front formed by normals to the family of the hyperbolas is curved. Show that the radius of curvature can be determined approximately by $R(z) \approx w(z)/[dw(z)/dz] = z(1 + z_F^2/z^2)$. Draw $w(z)$ and show that the focal points for the right-hand phase front and the left-hand phase front coincide when $z = z_F$. When the mirrors have radius of curvature $R(z_F) = 2z_F$ and are placed at a distance of $d = 2z_F$ apart, the configuration is confocal. Determine modes inside an optical cavity made of confocal mirrors. Are the modes with the confocal configuration stable?

Problem P5.12

Consider wave diffraction by a wedge as shown in Figure P5.12. The geometrical theory of diffraction (GTD) stipulates that the electric field takes the form

$$\overline{E} = \overline{E}^i + \overline{E}^r + \overline{E}^d$$

The diffracted field is

$$\overline{E}^d(s) = \overline{E}^i(0) \cdot \overline{\overline{D}}(\hat{s}, \hat{s}^i) A(s) e^{iks}$$

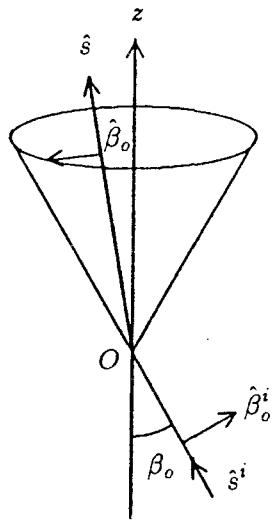


Figure P5.12b

where $\hat{s}$ is the direction of the diffracted ray, $\hat{s}^i$ the direction of the incident ray, and s the optical path length of the diffracted ray from the edge. The function $A(s) = l/\sqrt{s}$ for plane, cyclindrical, and conical wave incidence, and $A(s) = (s^i/s(s^i + s))^{1/2}$ for spherical wave incidence. The dyadic diffraction coefficient

$$\overline{\overline{D}}(\hat{s}, \hat{s}^i) = -\hat{\beta}_o^i \hat{\beta}_o D_v - \hat{\phi}^i \hat{\phi} D_h$$

where $\hat{\beta}_o^i = \hat{s}^i \times \hat{\phi}^i$ and $\hat{\beta}_o = \hat{s} \times \hat{\phi}$ as shown in Figure P5.12b. The first term corresponds to a TM wave and the second term to a TE wave. If the field points are not close to the shadow or reflection boundary, the scalar diffraction coefficients are

$$D_{v,h}(\phi, \phi^i, \beta_o) = \frac{e^{i\pi/4} \sin\frac{\pi}{n}}{n\sqrt{2\pi k} \sin\beta_o} \left[\frac{1}{\cos\frac{\pi}{n} - \cos\frac{\phi - \phi^i}{n}} \mp \frac{1}{\cos\frac{\pi}{n} - \cos\frac{\phi + \phi^i}{n}} \right]$$

where the upper sign is for v and the lower sign for h.

(a) Show that at $\phi = 0$ and $\phi = n\pi$, the tangential component of $\overline{E}^d$ satisfies the boundary condition.

(b) Show that at the shadow boundary ($\phi = \pi + \phi_i$) and at the reflection boundary ($\phi = \pi - \phi^i$), the GTD breaks down.

Problem P5.13

Find a stationary formula for cutoff frequencies in a cylindrical metallic waveguide similar to the one for resonator cavities.

Problem P5.14

Derive a stationary formula for the resonant frequency of a resonator cavity in terms of the magnetic field $\overline{H}$. Show that it is

$$k^2 = \frac{\iiint dV\,(\nabla \times \overline{H})^2}{\iiint dV\,|\overline{H}|^2}$$

and that $\overline{H}$ need not satisfy any boundary conditions. Repeat the example for a circular cavity as given in the text by assuming approximate magnetic field distributions.

Problem P5.15

Show that, if the inside of a waveguide is filled with several different isotropic media, the stationary formula for the cutoff frequency in terms of the $\overline{E}$ field is

$$\omega_c^2 = \frac{\iint dS\,\mu^{-1}(\nabla \times \overline{E})^2}{\iint dS\,\epsilon E^2} + \frac{2\int dl\,\hat{n} \cdot [\mu^{-1}(\nabla \times \overline{E}) \times \overline{E}]}{\iint dS\,\epsilon E^2}$$

where $\hat{n}$ is the outward-pointing unit vector normal to the waveguide walls.

Problem P5.16

Use the stationary formula as derived in Problem P5.15 to solve for the cutoff frequency of a rectangular waveguide filled with two different dielectric media. Let the waveguide dimension a be along $\hat{x}$ and b along $\hat{y}$. The dielectric ϵ_1 fills the space from $x = 0$ to $x = a/2$, and the dielectric ϵ_2 fills the space from $x = a/2$ to $x = a$. Assuming an electric field of $\overline{E} = \hat{y}\sin(\pi x/a)$, show that the cutoff frequency is

$$\omega_c \approx \frac{\pi}{a\sqrt{\mu(\epsilon_1 + \epsilon_2)/2}}$$

Problem P5.17

Show that the echo area for backscattering resulting from a linearly polarized field is

$$\overline{E}^i = \hat{z} E_0 e^{-ikz}$$

$$A_e = \lim_{r\to\infty} 4\pi r^2 \left|\frac{\overline{E}_s}{\overline{E}^i}\right|^2 = \pi \left|\frac{\eta}{\lambda} \frac{\left(\iint J_z^a e^{-ikz} ds\right)^2}{\iint \overline{E}^a \cdot \overline{J}^a ds}\right|^2$$

Problem P5.18

A plane wave $\overline{E}^i = \hat{z} E_0 e^{ikz}$ is normally incident upon a conducting wire of length $L = \lambda/2$ along the z axis. Assume that the current on the wire is $I^a = \cos kz$. Using the stationary formula in Problem P5.17 and the fact that $< a, a >= 73$, show that the echo area for backscattering is $A_e \to 0.86\lambda^2$.

Problem P5.19

Which of the following media are reciprocal? For the nonreciprocal ones, what are their complementary media?

(a) A biaxial medium.
(b) A moving biaxial medium.
(c) A biisotropic medium with a real χ.
(d) A biisotropic medium with an imaginary χ.
(e) A ferrite in a dc magnetic field.

Problem P5.20

For a cavity filled with several different isotropic media, show that the stationary formula with an assumed electric field is given by

$$\omega_r^2 = \frac{\iiint_V dV \mu^{-1} (\nabla \times \overline{E})^2 + 2 \iint_S dS\, \hat{n} \cdot \{\mu^{-1} (\nabla \times \overline{E}) \times \overline{E}\}}{\iiint_V dV\, \epsilon E^2}$$

where $\hat{n}$ is the outward-pointing unit vector normal to the cavity walls. Consider a microwave oven, which has dimensions $h = 25$ cm, $w = 40$

cm, $d = 40$ cm. Consider a dielectric of rectangular shape with permittivity $\epsilon = 4\epsilon_o$ and dimensions 15 cm$(h) \times 30$ cm$(w) \times 20$ cm(d). Assume the electric field $\overline{E} = \hat{x}\sin(\pi y/w)\sin(\pi z/d)$ and find the percentage of variation in resonant frequency.

Problem P5.21

Consider a lens formed by rotating a hyperbola around its axis. Show that the field amplitude distribution at the aperture plane is

$$\frac{E(\rho)}{E(0)} = \frac{1}{n-1}\sqrt{\frac{(n\cos\phi - 1)^3}{n - \cos\phi}}$$

where we assume the lens is illuminated by an isotropic source placed at the focal point to the left of the lens.

Problem P5.22

Let the gain function for the feed of a paraboloid reflector antenna be

$$\begin{aligned} G(\theta') &= 6\cos^2\theta' \qquad 0 \le \theta' \le \pi/2 \\ &= 0 \qquad\qquad\quad \pi/2 \le \theta' \end{aligned}$$

Show that

$$\int G(\theta')\, d\Omega = 4\pi$$

Show that with this gain function the optimum aperture angle $\theta_0 \approx 66°$.

Problem P5.23

The gain of a reflector antenna is affected by phase errors of the feed. Let the field intensity pattern of a paraboloidal reflector antenna feed have the form

$$\overline{E} = \hat{e}_0 \left[2\eta \frac{P_t}{4\pi} G(\theta)\right]^{1/2} e^{j\frac{\omega}{c}\delta(\theta)}$$

where $\delta(\theta)$ represents the phase error of the feed pattern. Show that the gain factor along the reflector axis is given by

$$g = \cot^2\left(\frac{\theta}{2}\right)\left\{\left[\int_0^{\theta_0} d\theta' \,[G(\theta')]^{1/2}\cos\left[\frac{2\pi\delta(\theta')}{\lambda}\right]\tan\frac{\theta'}{2}\right]^2 + \left[\int_0^{\theta_0} d\theta' \,[G(\theta')]^{1/2}\sin\left[\frac{2\pi\delta(\theta')}{\lambda}\right]\tan\frac{\theta'}{2}\right]^2\right\}$$

Problem P5.24

Follow the steps below to obtain the dyadic Green's function for a perfectly conducting half-space. The equation to be satisfied by $\overline{\overline{G}}(\bar{r}, \bar{r}')$ is

$$\nabla \times \nabla \times \overline{\overline{G}}(\bar{r}, \bar{r}') - k^2 \overline{\overline{G}}(\bar{r}, \bar{r}') = \overline{\overline{I}}\delta(\bar{r} - \bar{r}')$$

and the boundary condition is

$$0 = \hat{z} \times \overline{E}|_{z=0} = i\omega\mu \iiint_{V'} dV' \, [\hat{z} \times \overline{\overline{G}}(\bar{r}, \bar{r}')]_{z=0} \cdot \overline{J}(\bar{r}')$$

which must be satisfied for all $\overline{J}(\bar{r})$. Consequently,

$$\hat{z} \times G(\bar{r}, \bar{r}')\big|_{z=0} = 0$$

where $\hat{z}$ is the unit surface normal. Given any Hertzian dipole with arbitrary orientation, decompose it into two components — one parallel to the boundary, and the other perpendicular to the boundary. Using the image theory separately for the parallel and perpendicular components of the arbitrary Hertzian dipole, show that

$$\overline{\overline{G}}(\bar{r}, \bar{r}') = \overline{\overline{G}}_0(\bar{r} - \bar{r}') - \overline{\overline{G}}_0(\bar{r} - \bar{r}' \cdot (\overline{\overline{I}} - 2\hat{z}\hat{z})) \cdot (\overline{\overline{I}} - 2\hat{z}\hat{z})$$

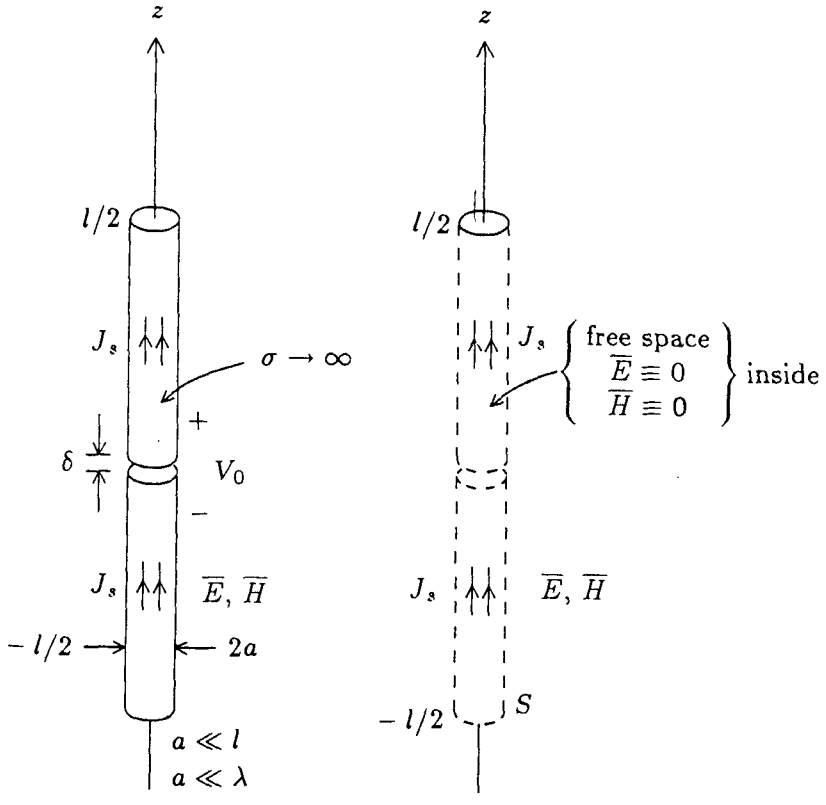
where $\overline{\overline{G}}_0$ is the free-space dyadic Green's function, and the origin is assumed to be on the surface, without loss of generality. The free-space dyadic Green's function is invariant under a translation of both the observation and the source points,

$$\overline{\overline{G}}_0(\bar{r} + \bar{r}_0, \bar{r}' + \bar{r}_0) = \overline{\overline{G}}_0(\bar{r}, \bar{r}')$$

Show that this property does not hold for the dyadic Green's function for the conducting half-space, unless $\bar{r}_0 \cdot \hat{z} = 0$. For what physical reason is this true?

Problem P5.25

Following the steps below, obtain Hallen's integral equation for a thin linear dipole antenna, and show that the current on the antenna is approximately sinusoidal.



Original problem Equivalent problem

Figure P5.25

First set up an equivalent problem in which $\overline{E} = 0$ and $\overline{H} = 0$ inside the surface S, which is identical to the surface of the conductors of the antenna and $\overline{E}$ and $\overline{H}$ are the same as in the original problem outside S. There are no conducting bodies in the equivalent problem. To accomodate the discontinuity of $\overline{H}_t$ across S, we must impress the surface current $\overline{J}_s$ on the imaginary surface S. This surface current is identical to the induced surface current in the original problem. Argue that the electric field along the z axis approaches

$$\overline{E} \to \hat{z} V_0 \delta(z) \qquad \text{for} \qquad |a| < l/2$$

as the gap becomes very small. Show that the electric field on the z axis is given by

$$E_z(z)|_{\rho=0} = V_0\delta(z) = i\omega\mu\hat{z}\cdot\left(\overline{\overline{I}} + \frac{\nabla\nabla}{k^2}\right)\int_{-l/2}^{l/2} dz' \cdot \int_0^{2\pi} a\,d\phi' \frac{e^{ik\sqrt{(z-z')^2+a^2}}}{4\pi\sqrt{(z-z')^2+a^2}}\,\overline{J}_s(z') \quad \text{for} \quad |z| < l/2$$

The contribution of the current at the caps of the cylinder is ignored to obtain this last result. Discuss why this is a good approximation if $a \ll \lambda$ and $a \ll l$.

Argue for the following assumptions:

(i) $\overline{J}_s(\overline{r}')$ is along the z axis.
(ii) $\overline{J}_s$ has no ϕ variation.

It follows from (i) and (ii) that $I(z') = 2\pi a J_s(z')$. With these simplifications, show that

$$E_z = \frac{i\omega\mu}{k^2}\left(k^2 + \frac{d^2}{dz^2}\right)\int_{-l/2}^{l/2} dz' \frac{e^{ik\sqrt{(z-z')^2+a^2}}}{4\pi\sqrt{(z-z')^2+a^2}} I(z') = V_0\,\delta(z)$$

Solve this differential equation to obtain

$$\frac{i\omega\mu}{k^2}\int_{-\pi/2}^{\pi/2} dz' \frac{e^{ik\sqrt{(z-z')^2+a^2}}}{4\pi\sqrt{(z-z')^2+a^2}} I(z') = V_0[g(z) + C_1\cos kz + C_2\sin kz]$$

where $g(z)$ is the scalar Green's function of the one-dimensional Helmholtz equation,

$$\left(\frac{d^2}{dz^2} + k^2\right) g(z) = \delta(z)$$

Show that

$$g(z) = \frac{1}{2ik} e^{ik|z|}$$

For the case of symmetric excitation, $I(z) = I(-z)$. Show that $C_2 = 0$ for this case. You have now obtained Hallen's integral equation

$$\int_{-l/2}^{l/2} dz'\; K(z, z')I(z') = \frac{4\pi k^2}{i\omega\mu} V_0\left(\frac{e^{ik|z|}}{2ik} + C_z\cos kz\right)$$

where

$$K(z, z') = \frac{e^{ik\sqrt{(z-z')^2+a^2}}}{\sqrt{(z-z')^2+a^2}}$$

Make a realistic sketch of $|K(z, z')|$ for fixed z' and $|z| < l/2$. Observe that $|K(z, z')|$ is highly peaked at $z = z'$ if $a \ll l$. In view of this fact, we write

$$I(z)\int_{-l/2}^{l/2} dz'\, K(z, z') + \int_{-l/2}^{l/2} dz'\, K(z, z')\left[I(z') - I(z)\right] = \frac{4\pi k^2}{i\omega\mu} V_0 \left(\frac{e^{ik|z|}}{2\pi ik} + C_1 \cos kz\right)$$

where the second integral is much smaller than the first one. Why?

Sketch for $|z| < l/2$,

$$f(z) = \int_{-l/2}^{l/2} dz' K(z, z')$$

Show that it is a very flat function of z, except at the ends. Observe that $\lim_{z\to\pm l/2} f(z) \neq 0$. It is a good approximation to assume that $f(z)$ is constant. $f(z)I(z) \simeq f(0)I(z)$. Evaluate C_1, imposing the condition that $I(\pm l/2) = 0$. Express $I(z)$ in the following form:

$$I(z) \simeq \frac{2kV_0}{i\omega\mu} \frac{\sin(k|z| - kl/2)}{f(0)\cos(kl/2)}$$

Problem P5.26

Consider a short-circuited, parallel-plate waveguide with length l in the $\hat{z}$ direction and separation d in the $\hat{x}$ direction. Find the zeroth-, first-, and second-order electric and magnetic fields. Show that the sum of all orders of the quasi-static solution is equal to the full-wave solution for the TEM mode. Assume that the driving current at $z = 0$ is $I(t) = I_0 \cos \omega t$ and that all fields are functions of t and z only.

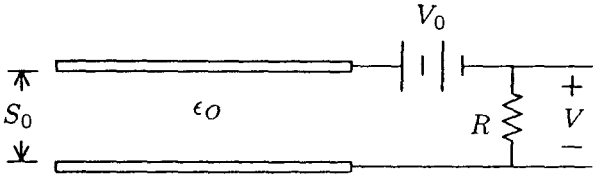


Figure P5.27

Problem P5.27

In the electroquasistatic field, the first-order current is obtained from the time derivative of the zeroth-order charge distribution. For a capacitor carrying charge $Q = CV$, the first-order current is

$$I = \frac{dQ}{dt} = C\frac{dV}{dt} + V\frac{dC}{dt}$$

Electroquasistatic transducers are devices that use the second term by mechanically vibrating a capacitor to generate electric current.

In a condenser microphone, a metal-plate diaphragm is tightly stretched but still capable of movement in response to sound transients. It is often made of a plastic film coated with an extremely fine, thin covering of gold to make it conductive. The diaphragm forms one plate of a capacitor with a dielectric plastic backing facing the fixed back-plate electrode. The equivalent circuit for a condenser microphone is shown in Figure P5.27.

For a time-harmonic sound pressure creating a diaphragm movement of $s(t) = S_1 \cos \omega t$, show that the differential equation for $v(t)$ is

$$RC_0\frac{dv(t)}{dt} + v(t) = RC_0V_0\frac{d}{dt}\left[\frac{s(t)}{S_0}\right]$$

where $C_0 = \epsilon_o A/S_0$ and A is the area of the diaphragm. Let $V_0 = 50$ volts, $R = 10^7$ ohms, $A = 8\ \text{cm}^2$, $f = 1$ kHz, $S_0 = 25\ \mu\text{m}$, and $S_1 = 1\ \mu\text{m}$. Find the magnitude of the output voltage. Show that the induced voltage is directly proportional to the diaphragm vibration amplitude S_1 when $\omega RC_0 \gg 1$.

Problem P5.28

Consider a plane wave normally incident on a conducting plate with area A. Assume that A is very large compared with the squared wavelength. Using the equivalence principle, approximate the source on the plate and show that the echo area is $A_e \approx \dfrac{k^2 A^2}{\pi}$.

Problem P5.29

(a) Consider a surface current sheet with current density given by

$$\overline{J} = \hat{x} J_s \delta(z)$$

Determine the time-average power per unit area generated by the current sheet and the time-average power per unit area carried by the plane wave in the half-space $z > 0$

(b) For a surface current sheet with the current density given by

$$\overline{J} = \hat{x} J_s e^{-j\beta y} \delta(z)$$

determine the amplitudes of the waves generated and the associated time-average power densities. What happens when $\beta = k = \omega\sqrt{\mu_o \epsilon_o}$?

(c) In the quasistatic limit, show that the magnetic field approaches that generated by a sinusoidal current sheet.

Problem P5.30

For a microstrip antenna consisting of a conducting patch printed on a dielectric slab with a ground plane conductor, the radiation field is produced by the surface electric currents on the conducting patch. In recent studies of its radiation characteristics, the equivalence principle has been used to calculate the fields produced by a ribbon-like magnetic surface current sheet based on the contention that the fields produced below the conducting patch closely resemble those in the cavity bounded by a magnetic wall surrounding the patch inside the substrate.

The printed patch is perfectly conducting and denoted by S_{a1} with surface normal $\hat{z}$. The substrate, signified as region 1, is a dielectric with permittivity ϵ_1, permeability μ, and thickness d. The free space is denoted as region 0 with permittivity ϵ and permeability μ. The substrate is placed on a ground plane. The normal projection

of the printed patch on the ground plane is denoted as S_{a2} with normal $-\hat{z}$. The ribbon-like side wall bounding S_{a1} and S_{a2} is called S_b.

We consider the cavity space bounded by S_b and $S_a = S_{a1} + S_{a2}$ and assume that S_b is a perfect magnetic conductor. Let the electric and magnetic fields inside the cavity be denoted by $\overline{E}_c$ and $\overline{H}_c$. The boundary conditions state that

$$\hat{n} \times \overline{E}_c = 0 \qquad \text{on } S_a$$
$$\hat{n} \times \overline{H}_c = 0 \qquad \text{on } S_b$$

The purpose is to show that the field outside the cavity produced by the electric current sheet on S_{a1} is equal to that produced by an equivalent magnetic current sheet on S_b.

(a) The dyadic Green's function of the third kind [Tai, 1971] satisfies the equation

$$\nabla' \times \nabla' \times \overline{\overline{G}}_{10}^{(3)}(\overline{r}', \overline{r}) - k_1^2 \overline{\overline{G}}_{10}^{(3)}(\overline{r}', \overline{r}) = 0$$

for $\overline{r}'$ in region 1 and $\overline{r}$ in region 0 where $k_1^2 = \omega^2 \mu \epsilon_1$. The electric field in region 1 inside the cavity satisfies the equation

$$\nabla' \times \nabla' \times \overline{E}_c(\overline{r}') - k_1^2 \overline{E}_c(\overline{r}') = 0$$

Making use of vector Green's theorem, the Maxwell equation $\nabla \times \overline{E}_c(\overline{r}')$, and $\hat{n} \times \overline{\overline{G}}_{10}^{(3)} = 0$ on S_{a2}, show that

$$0 = i\omega\mu \int_{S_{a1}} dS' \hat{n} \times \overline{H}_c(\overline{r}') \cdot \overline{\overline{G}}_{10}^{(3)}(\overline{r}', \overline{r}) + \int_{S_b} dS' \hat{n} \times \overline{E}_c(\overline{r}') \cdot \nabla' \times \overline{\overline{G}}_{10}^{(3)}(\overline{r}', \overline{r})$$

(b) The symmetry relations for the dyadic Green's functions state that

$$[\overline{\overline{G}}_{10}^{(3)}(\overline{r}', \overline{r})]^t = \overline{\overline{G}}_{01}^{(3)}(\overline{r}, \overline{r}')$$
$$[\nabla' \times G_{10}^{(3)}(\overline{r}', \overline{r})]^t = \frac{\epsilon_1}{\epsilon} \nabla \times \overline{\overline{G}}_{01}^{(4)}(\overline{r}, \overline{r}')$$

where $\overline{\overline{G}}_{01}^{(4)}(\bar{r},\bar{r}')$ is the dyadic Green's function of the fourth kind which has the symmetry property

$$[\overline{\overline{G}}_{01}^{(4)}(\bar{r}',\bar{r})]^t = \frac{\epsilon}{\epsilon_1}\overline{\overline{G}}_{10}^{(4)}(\bar{r},\bar{r}')$$

Show that

$$\overline{E}_J(\bar{r}) = \overline{E}_M(\bar{r})$$

where

$$\overline{E}_J(\bar{r}) \equiv -i\omega\mu \int_{S_{a1}} dS' \, \overline{\overline{G}}_{01}^{(3)}(\bar{r},\bar{r}') \cdot [\hat{n} \times \overline{H}_c(\bar{r}')]$$

is the electric field due to the equivalent electric current sheet $\hat{n} \times \overline{H}_c(\bar{r}')$ on S_{a1} and

$$\overline{E}_M(\bar{r}) \equiv \frac{\epsilon_1}{\epsilon} \int_{S_b} dS' \, \nabla \times \overline{\overline{G}}_{01}^{(4)}(\bar{r},\bar{r}') \cdot [\hat{n} \times \overline{E}_c(\bar{r}')]$$

is the electric field due to the equivalent magnetic current sheet $\hat{n} \times \overline{E}_c(\bar{r}')$ on S_b.

Problem P5.31

What are the expectation values for the field operator $\overline{D}$ with a given energy state $|n>$? What are the expectation values for $\overline{D}^2$?

Problem P5.32

Let A, B, and C be Hermitian operators, with $[A,B] = iC$. Define the mean-square deviation by $(\Delta A)^2 = <\psi|(A - <A>)^2|\psi>$ and $(\Delta B)^2 = <\psi|(B - <B>)^2|\psi>$ with respect to the state function $|\psi>$, where $<A> = <\psi|A|\psi>$ and $<B> = <\psi|B|\psi>$ are expectation values of A and B. Using the Schwartz inequality $<\cdot\phi|\phi><x|x> \geq |<\phi|x>|^2$, show that

$$(\Delta A)^2(\Delta B)^2 \geq \frac{1}{4}|<\psi|C|\psi>|^2$$

when $C = \hbar$, the commutation relation for A and B implies the uncertainty relation $\Delta A \Delta B \geq \hbar/2$.

Problem P5.33

The coherent state is also a minimum uncertainty state. Show that with the coherent state $(\Delta A)\cdot(\Delta D) = \hbar/2$.

Problem P5.34

Find eigenvalues and eigenvectors for $a + a^+$ and $a - a^+$.

Problem P5.35

Prove that $e^{xa}e^{xa^+} = e^{xa^+}e^{xa}e^{x^2}$ if $[a, a^+] = 1$.

Problem P5.36

What are the commutation relations for field vectors at different space–time points? Interpret these relations, and use them to derive Maxwell's equations from the postulated Hamiltonian.

Problem P5.37

Given $[a, a^+] = 1$, compute $[a, (a^+)^n]$ and $[a, e^{a^+}]$.

Problem P5.38

The Hamiltonian of a harmonic oscillator is given by

$$\mathcal{H} = \frac{1}{2m}(p^2 + m^2\omega^2 x^2)$$

Define $a = (m\omega x + ip)/\sqrt{2m\hbar\omega}$, $a^+ = (m\omega x - ip)/\sqrt{2m\hbar\omega}$. Given $[x, p] = i\hbar$, what are the Hamiltonian and commutation relations for a and a^+? Show that $(\Delta p)^2 = \frac{m\hbar\omega}{2}(2n+1)$ and $(\Delta x)^2 = \frac{\hbar}{2m\omega}(2n+1)^2$ by using the energy states of the Hamiltonian.

Problem P5.39

Show that the unitary transformation U which transforms from the Heisenberg picture to the Schrödinger picture is governed by the differential equation

$$i\hbar\frac{dU}{dt} = \mathcal{H}_s U$$

where $\mathcal{H}_s$ is the Hamiltonian in the Schrödinger picture.

Problem P5.40

The Schrödinger equation is

$$i\hbar\frac{d}{dt}|\psi_s> = \mathcal{H}_s|\psi_s>$$

Write a Hamiltonian in the Schrödinger picture as

$$\mathcal{H}_s = \mathcal{H}_{os} + \mathcal{H}_{is}$$

where $\mathcal{H}_{os}$ is time-independent and $\mathcal{H}_{is}$ is called the interaction Hamiltonian. Transform to the interaction picture by a unitary transformation V such that $\mathcal{H}_{ii} = V\mathcal{H}_{si}V^+$ and

$$i\hbar\frac{d}{dt}|\psi_i> = \mathcal{H}_{ii}|\psi_i>$$

where $\mathcal{H}_{ii}$ is the interaction Hamiltonian in the interaction picture. Show that the differential equation is equivalent to the following integral equation:

$$|\psi_i(t)> = |\psi_i(0) + \frac{1}{i\hbar}\int_0^t d\tau\, \mathcal{H}_{ii}(\tau)\,|\,\psi_i(\tau)>$$

For first-order perturbation, $|\psi_i(t)>$ is approximated by $|\psi_i(0)>$. Thus

$$|\psi_i(t)> = 1 + \frac{1}{i\hbar}\int_0^t d\tau\, \mathcal{H}_{ii}(\tau)\,|\,\psi_i(0)>$$

For second-order perturbation, this solution is substituted into the integral equation, and we have

$$|\,\psi_i(t)> = \left\{1 + \frac{1}{i\hbar}\int_0^t d\tau\, \mathcal{H}_{ii}(\tau)\left[1 + \frac{1}{i\hbar}\int_0^t d\tau'\, \mathcal{H}_{ii}(\tau')\right]\right\}|\,\psi_i(0)>$$

The process can be continued. The transition from state $|\psi_i(0)>$ to state $|\psi_i(t)>$ is governed by a scattering operator S, $|\psi_i(t)> = S\,|\,\psi_i(0)>$. Show that the transition amplitude between an initial state $|E_i>$ and a final state $|E_f>$ under the first order perturbation theory is

$$\begin{aligned} S_{fi} &= < E_f|S|\,E_i> \\ &= 2ie^{i(E_f-E_i)t/2\hbar} < E_f|\mathcal{H}_{ii}|\,E_i> \frac{\sin[(E_f-E_i)t/2\hbar]}{E_f-E_i} \end{aligned}$$

where $|E_f>$ and $|E_i>$ are the energy eigenvectors of the unperturbed Hamiltonian $\mathcal{H}_{os}$.

VI

SCATTERING

6.1 Scattering by Spheres
a. Rayleigh Scattering
b. Mie Scattering
6.2 Scattering by a Conducting Cylinder
a. Exact Solution
b. Watson Transformation
c. Creeping Waves
6.3 Scattering by Periodic Rough Surfaces
a. Scattering by Periodic Corrugated Conducting Surfaces
b. Scattering by Periodic Dielectric Surfaces
6.4 Scattering by Periodic Media
a. First-Order Coupled-Mode Equations
b. Reflection and Transmission by Periodically-Modulated Slab
c. Far-Field Diffraction of a Gaussian Beam
6.5 Scattering by Random Media
a. Dyadic Green's Function for Layered Media
b. Scattering by a Half-Space Random Medium
6.6 Scattering by Random Rough Surfaces
a. Kirchhoff Approximation
b. Geometrical Optics Solution
c. Small Perturbation Method
6.7 Effective Permittivity for a Volume Scattering Medium
a. Random Discrete Scatterers
b. Effective Permittivity for a Continuous Random Medium
Problems

6.1 Scattering by Spheres

a. Rayleigh Scattering

Rayleigh scattering characterizes the scattering of electromagnetic waves by particles much smaller than a wavelength. Consider a spherical particle with permittivity ϵ_s, permeability μ_s, and radius a, at the origin of a coordinate system [Fig. 6.1.1]. A plane wave polarized in the $\hat{z}$ direction is incident upon the particle, $\overline{E} = \hat{z}E_0 e^{ikz}$. Because the particle is very small, the scattered field is essentially that resulting from a point source. The $\hat{z}$ directed electric field induces a dipole moment, and the particle re-radiates as a dipole antenna. The solution takes the form

$$\overline{E} = \frac{-i\omega\mu Il\, e^{ikr}}{4\pi r}\left\{\hat{r}\left[\left(\frac{i}{kr}\right)^2 + \frac{i}{kr}\right]2\cos\theta + \hat{\theta}\left[\left(\frac{i}{kr}\right)^2 + \frac{i}{kr} + 1\right]\sin\theta\right\} \tag{1a}$$

$$\overline{H} = \hat{\phi}\frac{-ikIl\, e^{ikr}}{4\pi r}\left(\frac{i}{kr} + 1\right)\sin\theta \tag{1b}$$

The dipole moment Il will be determined by E_0 and ϵ_s.

Very close to the origin, $kr \ll 1$. This also corresponds to the static limit when the frequency is very low since $k = \omega/c$. The dipole solution is electric in nature and the magnetic field vanishes in the limit, since $|\,\overline{H}\,| \sim Il$ and $|\,\overline{E}\,| \sim Il/k$ while Il is proportional to ω. The electric field in the static limit is

$$\overline{E} \approx \frac{i\omega\mu Il}{4\pi r}\frac{1}{(kr)^2}(\hat{r}\,2\cos\theta + \hat{\theta}\sin\theta)$$

$$= (\hat{r}\,2\cos\theta + \hat{\theta}\sin\theta)\left(\frac{a}{r}\right)^3 E_s \tag{2a}$$

where

$$E_s = \frac{i\eta Il}{4\pi ka^3} \tag{2b}$$

and $\eta = \sqrt{\mu/\epsilon}$. This solution satisfies the Maxwell equations for static fields, that is, $\nabla \times \overline{E} = 0$ and $\nabla \cdot \overline{E} = 0$.

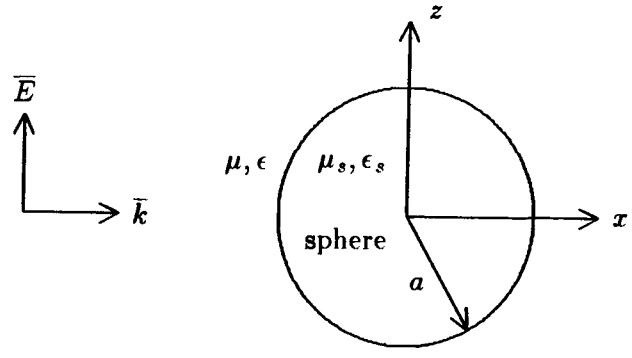


Figure 6.1.1 Rayleigh scattering by a small sphere.

We assume the field inside the sphere to be uniform and is in the same direction as the incident field:

$$\overline{E} = \hat{z} E_i = (\hat{r} \cos\theta - \hat{\theta} \sin\theta) E_i \qquad r \leq a$$

This solution also satisfies the Maxwell equations for static fields.

On the surface $r = a$, the boundary conditions require that the tangential $\overline{E}$ field and the normal $\overline{D}$ be continuous. We find

$$-E_0 + E_s = -E_i \tag{3a}$$

$$\epsilon E_0 + 2\epsilon E_s = \epsilon_s E_i \tag{3b}$$

In terms of the applied field E_0, we find from (3a) and (3b)

$$E_s = \frac{\epsilon_s - \epsilon}{\epsilon_s + 2\epsilon} E_0 \tag{4a}$$

$$E_i = \frac{3\epsilon}{\epsilon_s + 2\epsilon} E_0 \tag{4b}$$

From (2b) and (4a) we solve for Il and obtain

$$Il = -i4\pi k a^3 \sqrt{\frac{\epsilon}{\mu}} \left(\frac{\epsilon_s - \epsilon}{\epsilon_s + 2\epsilon} \right) E_0$$

When this is inserted in (1), we obtain the electromagnetic fields for Rayleigh scattering.

We shall now study the scattered field as $kr \gg 1$. Equation (1) gives

$$E_\theta = -\left(\frac{\epsilon_s - \epsilon}{\epsilon_s + 2\epsilon}\right) k^2 a^2 E_0 \frac{a}{r} e^{ikr} \sin\theta \tag{5a}$$

$$H_\phi = \sqrt{\frac{\epsilon}{\mu}} E_\theta \tag{5b}$$

The total scattered power from the sphere is

$$P_s = \frac{1}{2}\int_0^\pi r^2 \sin\theta \, d\theta \int_0^{2\pi} d\phi \, E_\theta H_\phi^* = \frac{4\pi}{3}\sqrt{\frac{\epsilon}{\mu}}\left(\frac{\epsilon_s - \epsilon}{\epsilon_s + 2\epsilon} k^2 a^3 E_0\right)^2 \tag{6}$$

The scattering cross section is calculated as

$$\Sigma_s = \frac{P_s}{\frac{1}{2}\sqrt{\frac{\epsilon}{\mu}} \mid E_0 \mid^2} = \frac{8\pi}{3}\left(\frac{\epsilon_s - \epsilon}{\epsilon_s + 2\epsilon}\right)^2 k^4 a^6 \tag{7}$$

Thus the total scattered power is proportional to the fourth power of the wavenumber; high-frequency waves are scattered more than lower ones. The scattered power is also proportional to the sixth power of the radius.

In the case of a perfectly conducting sphere, the electric field inside, $\overline{E}_i$, is identically zero. From (2*b*) and (3*a*) we find

$$Il = -i4\pi k a^3 \sqrt{\frac{\epsilon}{\mu}} E_0 \tag{8}$$

The boundary condition corresponding to (3*b*) for the normal $\overline{D}$ field can be used to find the surface charge densities ρ_s. Note that we can obtain (8) from (4) by letting $\epsilon_s \to \infty$. Since there are surface charges, their time variation will lead to a surface current, creating magnetic dipoles. The near field $\overline{H}$ for magnetic dipoles is the dual of that for the electric dipole in (2),

$$\overline{H} \sim \frac{ikKl}{4\pi r}\sqrt{\frac{\epsilon}{\mu}}\frac{1}{(kr)^2}\left(\hat{r}\, 2\cos\theta_y + \hat{\theta}_y \sin\theta_y\right) \tag{9}$$

where Kl is the magnetic moment of the dipole. Notice that for the incident $\overline{H}$ field in the $\hat{y}$ direction, the angle θ_y now refers to the

y axis as opposed to the z axis for the electric dipole. The boundary condition requires zero normal $\overline{B}$ field, and the discontinuity of the tangential $\overline{H}$ field gives rise to the surface current densities $\overline{J}_s$. We find

$$Kl = -i2\pi ka^3\sqrt{\frac{\mu}{\epsilon}}H_0 \tag{10}$$

where H_0 is the amplitude of the incident plane wave. Thus the scattered field corresponds to that of a magnetic dipole along the y axis.

We shall remember that the above analysis for Rayleigh scattering is valid only when the radius of the sphere is very small. For larger radii, the scattering process is called Mie scattering. In fact, as we shall show in the following, scattering of a plane wave by a sphere of arbitrary size with permittivity ϵ_s and permeability μ_s can be solved exactly in closed form.

b. Mie Scattering

The problem of a plane wave scattered by a sphere can be solved rigorously by matching boundary conditions. To facilitate the solution, we decompose the spherical waves into TM to $\hat{r}$ and TE to $\hat{r}$ components by introducing the Debye potentials π_e and π_m so that

$$\overline{A} = \overline{r}\pi_e \tag{11a}$$

$$\overline{H} = \nabla \times \overline{A} = \hat{\theta}\frac{1}{\sin\theta}\frac{\partial}{\partial\phi}\pi_e - \hat{\phi}\frac{\partial}{\partial\theta}\pi_e \tag{11b}$$

for TM to $\hat{r}$ waves, and

$$\overline{Z} = \overline{r}\pi_m \tag{12a}$$

$$\overline{E} = \nabla \times \overline{Z} = \hat{\theta}\frac{1}{\sin\theta}\frac{\partial}{\partial\phi}\pi_m - \hat{\phi}\frac{\partial}{\partial\theta}\pi_m \tag{12b}$$

for TE to $\hat{r}$ waves.

The Debye potentials π_e and π_m satisfy the Helmholtz equation in spherical coordinates

$$(\nabla^2 + k^2)\begin{Bmatrix}\pi_e\\ \pi_m\end{Bmatrix} = 0 \tag{13}$$

where

$$\nabla^2 = \frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \tag{14}$$

The solution to this equation is composed of superpositions of spherical Bessel functions, associated Legendre polynomials, and sinusoids. Using Maxwell's equations and (13), we find that the field components in spherical coordinates take the forms

$$E_r = \frac{i}{\omega\epsilon}\left(\frac{\partial^2}{\partial r^2} r\pi_e + k^2 r\pi_e\right) \tag{15a}$$

$$E_\theta = \frac{i}{\omega\epsilon}\frac{1}{r}\frac{\partial^2}{\partial r\partial\theta} r\pi_e + \frac{1}{\sin\theta}\frac{\partial}{\partial\phi}\pi_m \tag{15b}$$

$$E_\phi = \frac{i}{\omega\epsilon}\frac{1}{r\sin\theta}\frac{\partial^2}{\partial r\partial\phi} r\pi_e - \frac{\partial}{\partial\theta}\pi_m \tag{15c}$$

$$H_r = -\frac{i}{\omega\mu}\left(\frac{\partial^2}{\partial r^2} r\pi_m + k^2 r\pi_m\right) \tag{16a}$$

$$H_\theta = -\frac{i}{\omega\mu}\frac{1}{r}\frac{\partial^2}{\partial r\partial\theta} r\pi_m + \frac{1}{\sin\theta}\frac{\partial}{\partial\phi}\pi_e \tag{16b}$$

$$H_\phi = -\frac{i}{\omega\mu}\frac{1}{r\sin\theta}\frac{\partial^2}{\partial r\partial\phi} r\pi_m - \frac{\partial}{\partial\theta}\pi_e \tag{16c}$$

The total electromagnetic fields are now decomposed into TE and TM components and expressed in terms of the Debye potentials π_e and π_m.

Consider a sphere with radius a located at the origin of a coordinate system [Fig. 6.1.2]. The sphere has permittivity ϵ_s and permeability μ_s. A plane wave,

$$\overline{E} = \hat{x}E_0\, e^{ikz} = \hat{x}E_0\, e^{ikr\cos\theta}$$

$$\overline{H} = \hat{y}\sqrt{\frac{\epsilon}{\mu}}E_0\, e^{ikr\cos\theta}$$

is incident upon the sphere. Note that the direction of propagation of the plane wave is along $\hat{z}$. The coordinate system is different from that used for Rayleigh scattering, where the z axis is in the direction of the linearly polarized electric field.

To match boundary conditions at the sphere's surface, we expand the incident wave in terms of spherical harmonics by using the wave transformation:

$$e^{ikr\cos\theta} = \sum_{n=0}^{\infty}(-i)^{-n}(2n+1)j_n(kr)P_n(\cos\theta) \tag{17}$$

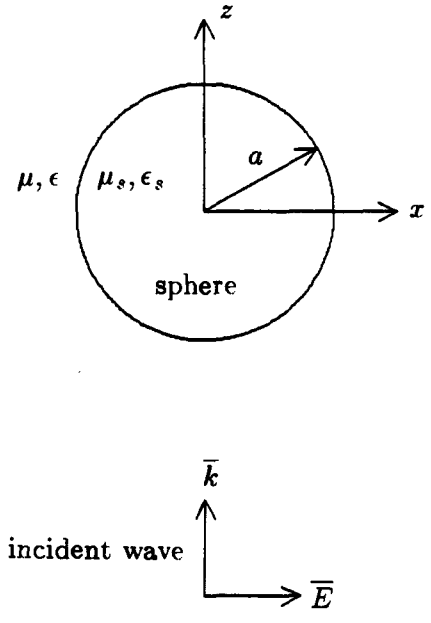


Figure 6.1.2 Mie scattering.

To determine the Debye potentials for the incident wave, we note that

$$E_r = E_0 \sin\theta \cos\phi \, e^{ikr\cos\theta}$$
$$= -\frac{iE_0 \cos\phi}{(kr)^2} \sum_{n=1}^{\infty} (-i)^{-n}(2n+1)\hat{J}_n(kr)P_n^1(\cos\theta)$$

where

$$\hat{J}_n(kr) = kr j_n(kr)$$

The summation now starts with $n = 1$ because $P_0^{(1)}(\cos\theta) = 0$. The potential π_e satisfies (15) and can be shown to be

$$\pi_e = -\frac{E_0 \cos\phi}{\omega\mu r} \sum_{n=1}^{\infty} \frac{(-i)^{-n}(2n+1)}{n(n+1)} \hat{J}_n(kr)P_n^1(\cos\theta) \qquad (18a)$$

By a dual process, the potential π_m is found to be

$$\pi_m = \frac{E_0 \sin\phi}{kr} \sum_{n=1}^{\infty} \frac{(-i)^{-n}(2n+1)}{n(n+1)} \hat{J}_n(kr)P_n^1(\cos\theta) \qquad (18b)$$

The scattered field can be characterized by Debye potentials:

$$\pi_e^s = -\frac{E_0 \cos\phi}{\omega\mu r} \sum_{n=1}^{\infty} a_n \hat{H}_n^{(1)}(kr) P_n^1(\cos\theta) \tag{19a}$$

$$\pi_m^s = \frac{E_0 \sin\phi}{kr} \sum_{n=1}^{\infty} b_n \hat{H}_n^{(1)}(kr) P_n^1(\cos\theta) \tag{19b}$$

where

$$\hat{H}_n^{(1)}(kr) = krh_n^{(1)}(kr)$$

The total field outside the sphere is equal to the sum of the incident and scattered fields.

The field inside the sphere can also be expressed in terms of the Debye potentials:

$$\pi_e^i = -\frac{E_0 \cos\phi}{\omega\mu_s r} \sum_{n=1}^{\infty} c_n \hat{J}_n(k_s r) P_n^1(\cos\theta) \tag{20a}$$

$$\pi_m^i = \frac{E_0 \sin\phi}{k_s r} \sum_{n=1}^{\infty} d_n \hat{J}_n(k_s r) P_n^1(\cos\theta) \tag{20b}$$

The boundary conditions at $r = a$ require that E_θ, E_ϕ, H_θ, and H_ϕ be continuous, with the result that four equations are solvable for the unknown coefficients a_n, b_n, c_n, and d_n. In view of (15) and (16), the coefficients are determined as

$$a_n = \frac{(-i)^{-n}(2n+1)}{n(n+1)} \cdot \frac{-\sqrt{\epsilon_s\mu}\,\hat{J}_n'(ka)\hat{J}_n(k_s a) + \sqrt{\epsilon\mu_s}\,\hat{J}_n(ka)\hat{J}_n'(k_s a)}{\sqrt{\epsilon_s\mu}\,\hat{H}_n^{(1)\prime}(ka)\hat{J}_n(k_s a) - \sqrt{\epsilon\mu_s}\,\hat{H}_n^{(1)}(ka)\hat{J}_n'(k_s a)} \tag{21a}$$

$$b_n = \frac{(-i)^{-n}(2n+1)}{n(n+1)} \cdot \frac{-\sqrt{\epsilon_s\mu}\,\hat{J}_n(ka)\hat{J}_n'(k_s a) + \sqrt{\epsilon\mu_s}\,\hat{J}_n'(ka)\hat{J}_n(k_s a)}{\sqrt{\epsilon_s\mu}\,\hat{H}_n^{(1)}(ka)\hat{J}_n'(k_s a) - \sqrt{\epsilon\mu_s}\,\hat{H}_n^{(1)\prime}(ka)\hat{J}_n(k_s a)} \tag{21b}$$

$$c_n = \frac{(-i)^{-n}(2n+1)}{n(n+1)}$$

$$\cdot \frac{i\sqrt{\epsilon_s\mu}}{\sqrt{\epsilon_s\mu}\,\hat{H}_n^{(1)\prime}(ka)\hat{J}_n(k_s a) - \sqrt{\epsilon\mu_s}\,\hat{H}_n^{(1)}(ka)\hat{J}_n'(k_s a)} \tag{21c}$$

$$d_n = \frac{(-i)^{-n}(2n+1)}{n(n+1)}$$

$$\cdot \frac{-i\sqrt{\epsilon\mu_s}}{\sqrt{\epsilon_s\mu}\,\hat{H}_n^{(1)}(ka)\hat{J}_n'(k_s a) - \sqrt{\epsilon\mu_s}\,\hat{H}_n^{(1)\prime}(ka)\hat{J}_n(k_s a)} \tag{21d}$$

In the case of small spheres, $ka \ll 1$ and $k_s a \ll 1$, only the $n = 1$ terms dominate, with $a_n \to -(ka)^3(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)$, and $b_n \to -(ka)^3(\mu_s - \mu)/(\mu_s + 2\mu)$. The results reduce to those of Rayleigh scattering. For the scattering of electromagnetic waves by spheres of finite radius, where the Rayleigh limit $ka \ll 1$ is not met, the phenomenon is known as *Mie scattering.*

The above results also reduce to the case of a perfectly conducting sphere. In the limit of $\epsilon_s \to \infty$ or $E_i = 0$ for a perfect conductor, the source-free Ampere's law $\nabla \times \overline{H} = -i\omega\epsilon_s\overline{E}$ gives a finite $\overline{H}$. From Faraday's law $\nabla \times \overline{E} = i\omega\overline{B}$, we see that a finite $\overline{B}$ field will give rise to a finite $\overline{E}$ field, violating the definition of a perfect conductor which requires $\overline{E}$ to be zero. Thus, $\overline{B}$ must be zero. However, there are no explicit mathematical requirements that $\overline{H}$ must be zero. If we let the permeability of the perfect conductor be μ_s such that $\overline{B} = \mu_s\overline{H}$, we see that $\mu_s = 0$. We can thus replace the definition of a zero $\overline{E}$ field for a perfect conductor by mathematically letting $\epsilon_s \to \infty$ and $\mu_s = 0$ when the general case of a medium with permittivity ϵ_s and permeability μ_s is solved. As a dual case, we can also define a perfect magnetic conductor by requiring $\overline{H}$ be identically zero, which leads to the mathematical limit of $\mu_s \to \infty$ and $\epsilon_s = 0$.

6.2 Scattering by a Conducting Cylinder

a. Exact Solution

Consider a plane wave incident upon a conducting cylinder [Fig. 6.2.1]. The incident wave is linearly polarized with electric vector $\overline{E}^i$ parallel to the axis of the cylinder. The incident $\overline{k}$ vector is perpendicular to the axis of the cylinder. In terms of cylindrical coordinates,

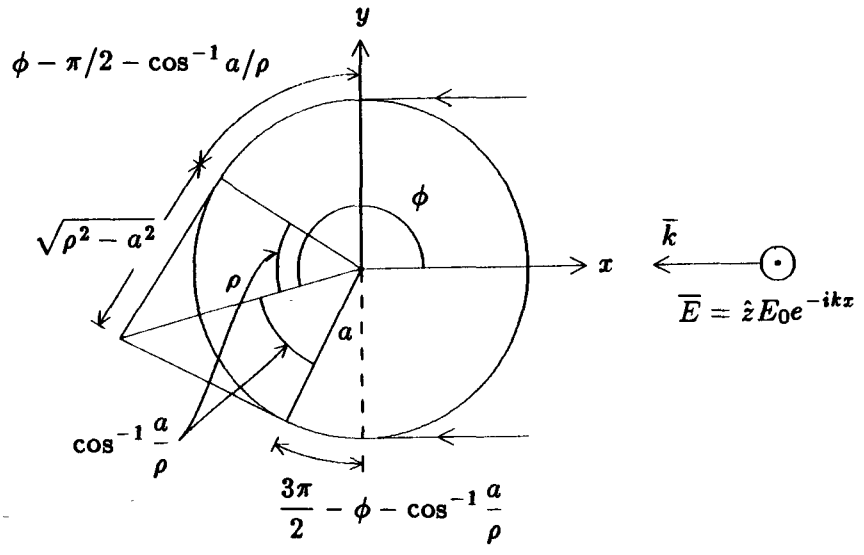


Figure 6.2.1 Scattering by a conducting cylinder.

we have

$$\bar{E}^i = \hat{z} E_0 e^{-ikx} = \hat{z} E_0 e^{-ik\rho \cos\phi} \tag{1}$$

To match boundary conditions at $\rho = a$, we transform the plane wave solution into a superposition of cylindrical waves satisfying the Helmholtz wave equation in cylindrical coordinates:

$$e^{-ik\rho\cos\phi} = \sum_{m=-\infty}^{\infty} a_m J_m(k\rho) e^{im\phi}$$

The constant a_m can be determined by using orthogonality relations for $e^{im\phi}$. We multiply both sides by $e^{-in\phi}$ and integrate over ϕ from 0 to 2π. In view of the integral representation for the Bessel function,

$$J_n(k\rho) = \frac{1}{2\pi} \int_0^{2\pi} d\phi\, e^{-ik\rho\cos\phi - in\phi + in\pi/2}$$

We obtain $a_n = e^{-in\pi/2}$ and

$$e^{-ik\rho\cos\phi} = \sum_{n=-\infty}^{\infty} J_n(k\rho) e^{in\phi - in\pi/2} \tag{2}$$

This expression is referred to as the *wave transformation*, which represents a plane wave in terms of cylindrical waves.

The scattered wave can also be expressed as a superposition of the cylindrical functions satisfying the Helmholtz wave equation. Expecting outgoing waves, we write the solution in terms of Hankel functions of the first kind. The sum of the incident wave and the scattered wave satisfies the boundary condition of a vanishing tangential electric field at $\rho = a$. We find the total solution to be

$$\overline{E} = \hat{z} E_0 \sum_{n=-\infty}^{\infty} \left[J_n(k\rho) - \frac{J_n(ka)}{H_n^{(1)}(ka)} H_n^{(1)}(k\rho) \right] e^{in\phi - in\pi/2} \tag{3}$$

The first summation term represents the incident wave; the second summation term, the scattered wave. At $\rho = a$, we find from (3) $\overline{E}(\rho = a) = 0$.

In the far-field zone, $k\rho \gg 1$, we make use of the asymptotic formula for $H_n^{(1)}(k\rho)$ and find that the scattered wave takes the form

$$\overline{E}_s \approx -\hat{z} E_0 \sum_{n=-\infty}^{\infty} \sqrt{\frac{2}{\pi k\rho}} \frac{J_n(ka)}{H_n^{(1)}(ka)} e^{ik\rho + in(\phi - \pi) - i\pi/4}$$

We can expand this result with respect to ka:

$$\overline{E}_s = \hat{z} i E_0 \sqrt{\frac{\pi}{2k\rho}} \left[\frac{1}{\ln ka} + (ka)^2 \cos\phi - \frac{(ka)^4}{8} \cos 2\phi + \cdots \right] e^{ik\rho - i\pi/4} \tag{4}$$

Observe that this series converges rapidly when the radius of the cylinder is small compared with the wavelength, $ka \ll 1$. The first term is angle-independent and signifies that the scattered wave caused by a thin wire is isotropic. We can show, however, that for an incident wave with magnetic field $\overline{H}$ parallel to the axis of the cylinder the scattered wave will no longer be isotropic and will be angle-dependent.

b. Watson Transformation

The series in (4) converges very slowly when the radius of the cylinder is not small compared to the wavelength. In this case, we use the Watson transformation to convert the solution into a rapidly convergent series. The Watson transformation relates a residue series

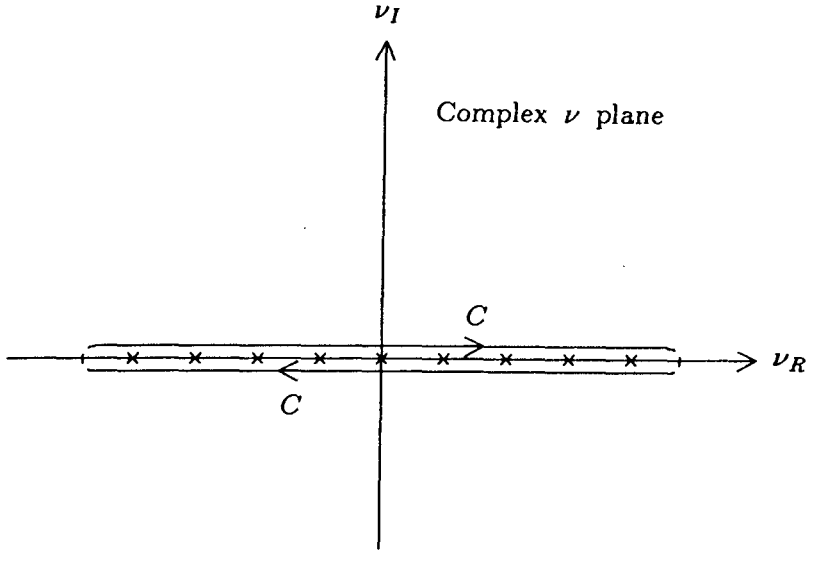


Figure 6.2.2 Watson transformation.

to a contour integration. Let the contour C in the complex ν plane be as depicted in Figure 6.2.2. We find

$$\sum_{n=-\infty}^{\infty} e^{in\phi} B_n = \frac{i}{2} \oint_C d\nu \frac{e^{i\nu(\phi-\pi)}}{\sin \nu\pi} B_\nu \tag{5}$$

under the assumption that B_ν has no singularities on the real axis. The singularities from $\sin \nu\pi$ are all first-order poles located on the real axis at $\nu = 0, \pm 1, \pm 2, \ldots$ Note that because of the direction of the contour, the contour integral is equal to $-2\pi i$ times the residues of the function $e^{i\nu(\phi-\pi)} B_\nu / \sin \nu\pi$, which are $e^{in\phi} B_n / \pi$ for all integer values of n.

To make use of the Watson transformation, we identify B_n according to (3):

$$\begin{aligned} B_n &= \frac{E_0}{H_n^{(1)}(ka)} \left[J_n(k\rho) H_n^{(1)}(ka) - J_n(ka) H_n^{(1)}(k\rho) \right] e^{-in\pi/2} \\ &= \frac{E_0}{2H_n^{(1)}(ka)} \left[H_n^{(2)}(k\rho) H_n^{(1)}(ka) - H_n^{(2)}(ka) H_n^{(1)}(k\rho) \right] e^{-in\pi/2} \end{aligned} \tag{6}$$

The singularities of B_n are caused by the zeros of $H_n^{(1)}(ka)$, which are illustrated in Figure 6.2.3.

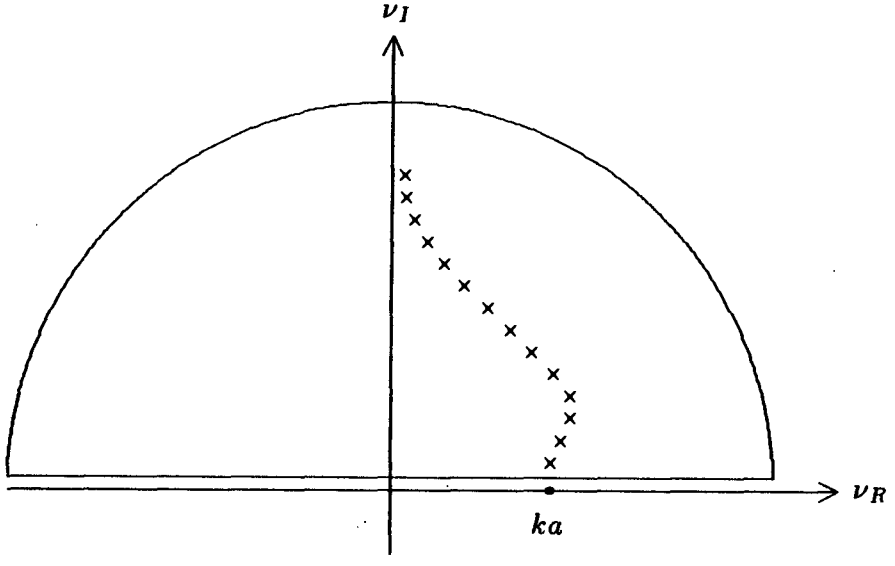


Figure 6.2.3 Integration over complex ν plane.

The contour integration on the right-hand side of (5) can be carried out by converting it into an integral from $-\infty$ to $+\infty$ and then closing the contour on the upper half-plane [Fig. 6.2.3]. From (5), we find

$$\sum_{n=-\infty}^{\infty} e^{in\phi} B_n = \lim_{\delta \to 0} \frac{i}{2} \left\{ \int_{-\infty+i\delta}^{\infty+i\delta} + \int_{\infty-i\delta}^{-\infty-i\delta} \right\} d\nu \; \frac{e^{i\nu(\phi-\pi)}}{\sin \nu\pi} B_\nu$$

$$= \frac{i}{2} \int_{-\infty}^{\infty} d\nu \, \frac{1}{\sin \nu\pi} \left\{ e^{i\nu(\phi-\pi)} B_\nu + e^{-i\nu(\phi-\pi)} B_{-\nu} \right\}$$

Using the relations

$$H_{-\nu}^{(1)}(\xi) = e^{i\nu\pi} H_\nu^{(1)}(\xi)$$
$$H_{-\nu}^{(2)}(\xi) = e^{-i\nu\pi} H_\nu^{(2)}(\xi)$$

we find $B_{-\nu} = B_\nu$. It can be shown that integration along the large semicircle makes no contribution. The entire contribution comes from the singularities of B_ν. Equation (5) yields

$$\overline{E} = \hat{z} i \int_{-\infty}^{\infty} d\nu \, \frac{\cos \nu(\phi - \pi)}{\sin \nu\pi} B_\nu$$

$$= \hat{z} \pi E_0 \sum_{n=1}^{\infty} \frac{H_{\nu_n}^{(2)}(ka)}{[\partial H_\nu^{(1)}(ka)/\partial \nu]_{\nu=\nu_n}} \frac{\cos \nu_n(\phi - \pi) e^{-i\nu_n \pi/2}}{\sin \nu_n \pi} H_{\nu_n}^{(1)}(k\rho) \quad (7)$$

where ν_n denotes zeros of $H_\nu^{(1)}(ka)$. Note that the first term of B_n in (6) does not contribute because $H_n^{(1)}(ka)$ in the numerator and the denominator cancel each other.

c. *Creeping Waves*

The series in (7) converges rapidly when $\pi/2 < \phi < 3\pi/2$. This is due to the fact that

$$\frac{\cos \nu_n(\phi - \pi)e^{-i\nu_n\pi/2}}{\sin \nu_n\pi} = \frac{-i\left[e^{i\nu_n(\phi-\pi/2)} + e^{i\nu_n(3\pi/2-\phi)}\right]}{1 - e^{i\nu_n 2\pi}} \tag{8a}$$

and ν_n takes positive imaginary values. The convergent range is in the shadow region of the cylinder. An interesting interpretation can be given to the terms involved. We use the asymptotic formula for $H_{\nu_n}^{(1)}(k\rho)$ when $k\rho > \nu_n \gg 1$:

$$H_{\nu_n}^{(1)}(k\rho) \sim \sqrt{\frac{2}{\pi(k^2\rho^2 - \nu_n^2)^{1/2}}}\, e^{i\left(\sqrt{k^2\rho^2-\nu_n^2} - \nu_n \cos^{-1}(\nu_n/k\rho) - \pi/4\right)} \tag{8b}$$

Note that the imaginary part of ν_n is positive and increases for increasing n. For the first few dominant terms we may approximate $\nu_n \sim ka$,

$$\sqrt{k^2\rho^2 - \nu_n^2} \approx k\sqrt{\rho^2 - a^2}$$

and

$$\cos^{-1}\frac{\nu_n}{k\rho} \approx \cos^{-1}\frac{a}{\rho}$$

In view of (8), the exponential dependence of (7) takes the form

$$e^{ik\sqrt{\rho^2-a^2}}\left\{e^{i\nu_n[\phi-(\pi/2)-\cos^{-1}(a/\rho)]} + e^{i\nu_n[3\pi/2-\phi-\cos^{-1}(a/\rho)]}\right\} \tag{9}$$

Now consider rays or photons incident upon the cylinder at tangent points, traveling along the surface, leaving the cylinder surface, and reaching the observation point at (ρ, ϕ) [Fig. 6.2.1]. The terms $(\phi - \pi/2 - \cos^{-1} a/\rho)$ and $(3\pi/2 - \phi - \cos^{-1} a/\rho)$ correspond to the paths of the two rays along the surface of the cylinder. The attenuation of the two rays is determined by the imaginary part of ν_n. The term $k(\rho^2 - a^2)^{1/2}$ corresponds to the path along which the rays travel after leaving the cylinder surface. The two rays recombine at ρ in the

shadow region. Because of this vivid picture, they are called *creeping waves.*

In the illuminated region, $-\pi/2 < \phi < \pi/2$, a useful formula can be obtained by noting that

$$e^{i\nu(\phi-\pi)} + e^{-i\nu(\phi-\pi)} = -i2e^{i\nu\phi}\sin\nu\pi + 2e^{i\nu\pi}\cos\nu\phi$$

When this relation is used, the contour integral in (5) becomes

$$\overline{E} = \hat{z}\left(\int_{-\infty}^{\infty} d\nu\, e^{i\nu\phi}B_\nu + i\int_{-\infty}^{\infty} d\nu\,\frac{\cos\nu\phi}{\sin\nu\pi}e^{i\nu\pi}B_\nu\right)$$

instead of (7). Upon closing the contour in the upper half-plane, the second integral is evaluated in terms of residues from $H_n^{(1)}(ka)$. Because the zeroes of $H_n^{(1)}(ka)$ have positive imaginary parts, this contribution will not be as significant as that of the first integral in the parentheses. The evaluation of the first integral is left as an exercise to the reader. The result is

$$\overline{E} \sim \hat{z}E_0\left[e^{-ik\rho\cos\phi} - \sqrt{\frac{a\cos\phi/2}{2\rho}}\,e^{ik(\rho-2a\cos\phi/2)}\right] \qquad (10)$$

This can be interpreted in terms of geometrical optics. The first term is the incident wave, and the second term corresponds to the ray reflected at the surface of the cylinder. Upon striking the surface at $\rho = a$, the incident ray has a phase factor $e^{-i(ka\cos\phi/2)}$. Upon reflection from the cylinder the ray reaches the observation point and gains another phase factor, $e^{-i(k\rho-ka\cos\phi/2)}$.

6.3 Scattering by Periodic Rough Surfaces

a. Scattering by Periodic Corrugated Conducting Surfaces

Consider a perfectly conducting surface that is corrugated periodically with rectangular grooves [Fig. 6.3.1] that are infinite in the $\hat{y}$ direction and have width w and depth d. The period of the corrugation is p. A plane wave with incident wave vector $\overline{k} = \hat{x}k_x - \hat{z}k_z$ is scattered by the rough surface. The scattered wave can be expanded

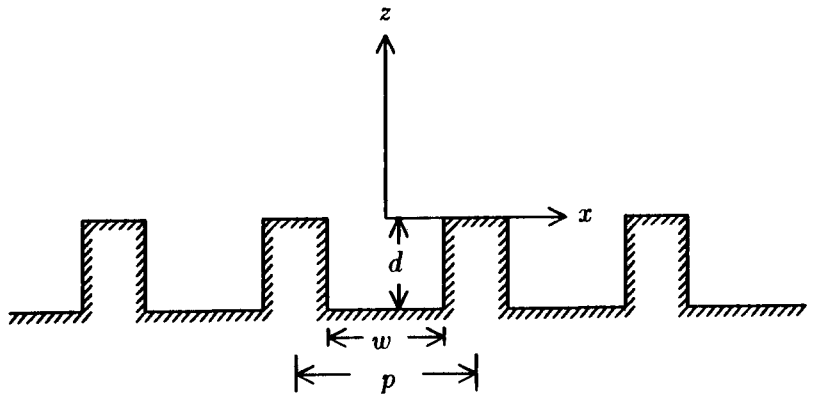


Figure 6.3.1 Scattering by a periodic corrugated surface.

in terms of Floquet modes that are characteristic waves for periodic structures. For an incident TM wave,

$$\overline{H} = \hat{y} H_0 \left\{ e^{ik_x x - ik_z z} + \sum_{n=-\infty}^{\infty} R_n \, e^{i(k_x + 2n\pi/p)x + ik_{zn} z} \right\} \qquad (1a)$$

where

$$k_{zn} = \sqrt{k^2 - (k_x + 2n\pi/p)^2} \qquad (1b)$$

The first term in $(1a)$ denotes the incident wave. The summation term represents a superposition of the Floquet modes. The electric field follows from the Maxwell equation

$$\overline{E} = \frac{i}{\omega\epsilon} \nabla \times \overline{H} \qquad (2)$$

We observe that for $k^2 < [k_x + (2n\pi/p)]^2$ the Floquet modes are essentially evanescent waves that decay exponentially in the $\hat{z}$ direction. The $n = 0$ mode corresponds to the specularly reflected wave.

Within the grooves the field can be expanded in terms of waveguide modes:

$$\overline{H} = \hat{y} H_0 \sum_{m=0}^{\infty} G_m \cos \frac{2m\pi}{w}(x + w/2) \frac{\cos k_{zm}^{(2)}(z + d)}{\cos k_{zm}^{(2)} d} \qquad z < 0 \quad (3a)$$

where

$$k_{zm}^{(2)} = \sqrt{k^2 - (2m\pi/w)^2} \tag{3b}$$

The electric field is derived from the magnetic field $\overline{H}$ by using the Maxwell equation

$$\begin{aligned}\overline{E} &= \frac{i}{\omega\epsilon}\nabla \times \overline{H} = \frac{i}{\omega\epsilon}\left\{\hat{z}\frac{\partial}{\partial x}H_y - \hat{x}\frac{\partial}{\partial z}H_y\right\} \\ &= \hat{x}\frac{ik_{zm}}{\omega\epsilon}H_0\sum_{m=0}^{\infty}G_m\cos\frac{m\pi}{w}(x+w/2)\frac{\sin k_{zm}^{(2)}(z+d)}{\cos k_{zm}^{(2)}d} \\ &\quad - \hat{z}\frac{im\pi}{\omega\epsilon w}H_0\sum_{m=0}^{\infty}G_m\sin\frac{m\pi}{w}(x+w/2)\frac{\cos k_{zm}^{(2)}(z+d)}{\cos k_{zm}^{(2)}d}\end{aligned} \tag{4}$$

It is straightforward to show that the fields in the grooves indeed satisfy the boundary conditions of tangential $\overline{E}$ fields vanishing at the perfectly conducting surfaces.

Consider the special case of normal incidence when $k_x = 0$. At $z = 0$, the boundary conditions require that the tangential magnetic field be continuous for $-w/2 \le x \le w/2$,

$$1 + \sum_{n=-\infty}^{\infty} R_n e^{i2n\pi x/p} = \sum_{m=0}^{\infty} G_m \cos\frac{m\pi}{w}(x+w/2) \tag{5a}$$

The tangential electric field is also continuous for $-w/2 \le x \le w/2$

$$1 - \sum_{n=-\infty}^{\infty}\frac{k_{zn}}{k_z}R_n e^{i2n\pi x/p} = -\sum_{m=0}^{\infty} i\frac{k_{zm}^{(2)}}{k_z}G_m\tan\left(k_{zm}^{(2)}d\right)\cos\frac{m\pi(x+w/2)}{w} \tag{5b}$$

The tangential electric field vanishes for $w/2 \le |x| \le p/2$,

$$1 - \sum_{n=-\infty}^{\infty}\frac{k_{zn}}{k_z}R_n e^{i2n\pi x/p} = 0 \tag{5c}$$

The task is to determine R_n and G_m from (5). We use the orthogonality properties for cosine functions by multiplying both sides of (5a) by $\cos[m\pi(x + w/2)/w]$ and integrating over the interval $-w/2 \le x \le w/2$. We obtain

$$[1+\delta_{m0}]\frac{w}{2}G_m = \int_{-w/2}^{w/2} dx\cos\frac{m\pi(x+w/2)}{w}\left[1 + \sum_{n=-\infty}^{\infty}R_n e^{i2n\pi x/p}\right] \tag{6}$$

We multiply (5*b*) by $e^{-i2n\pi x/p}$ and integrate over the interval $-p/2 \leq x \leq p/2$. By virtue of (5*c*), we know the right-hand side is nonzero only in the interval $-w/2 \leq x \leq w/2$. It follows that

$$p\left[\delta_{n0} - \frac{k_{zn}}{k_z} R_n\right] = -i \int_{-w/2}^{w/2} dx\, e^{-2n\pi x/p} \cdot \sum_{m=0}^{\infty} \frac{k_{zm}^{(2)}}{k_z} G_m \tan k_{zm}^{(2)} d \cos \frac{m\pi(x + w/2)}{w} \tag{7}$$

We define

$$P_{mn} = \int_{-w/2}^{w/2} dx\, e^{i2m\pi x/p} \cos \frac{n\pi(x + w/2)}{w} = \begin{cases} \dfrac{4m\pi/p}{(2m\pi/p)^2 - (n\pi/w)^2} \sin \dfrac{m\pi w}{p} & n\text{: even} \\ i \dfrac{4m\pi/p}{(2m\pi/p)^2 - (n\pi/w)^2} \cos \dfrac{m\pi w}{p} & n\text{: odd} \\ w\delta_{0n} & n = 0 \end{cases} \tag{8}$$

Equations (6) and (7) can be written as

$$[1 + \delta_{m0}] \frac{w}{2} G_m = P_{0m} + \sum_{n=-\infty}^{\infty} P_{nm} R_n \tag{9}$$

$$R_n = \left\{ \frac{k_z}{k_{zn}} \delta_{n0} + \sum_{m=0}^{\infty} \frac{i k_{zm}^{(2)}}{p k_{zn}} \tan\left(k_{zm}^{(2)} d\right) P_{nm}^* G_m \right\} \tag{10}$$

This represents a set of matrix equations to be solved for R_n. Substituting (10) into (9), we find

$$\sum_{l=0}^{\infty} \left[(1 + \delta_{l0}) \frac{w}{2} \delta_{ml} - i \tan\left(k_{zl}^{(2)} d\right) Q_{ml} \right] G_l = 2P_{0m} \tag{11}$$

where

$$Q_{ml} = \sum_{n=-\infty}^{\infty} \frac{k_z^{(2)} l}{p k_{zn}} P_{nm} P_{nl}^* = \begin{cases} \dfrac{1}{p} \left(P_{0m} P_{0l}^* + 2 \sum_{n=1}^{\infty} \dfrac{k_z^{(2)} l}{k_{zn}} P_{nm} P_{nl}^* \right) & m + l = \text{even} \\ 0 & m + l = \text{odd} \end{cases} \tag{12}$$

The mode amplitudes G_l are solved by straightforward matrix inversion. The number of groove modes needed to calculate the reflection coefficient R_n is determined by the width of the grooves w.

For sufficiently narrow grooves, $kw \ll 1$. We use the lowest mode amplitudes G_0 to calculate R_n. With $m = 0$, (6) and (7) become

$$G_0 = 1 + \sum_{n=-\infty}^{\infty} R_n \frac{p}{n\pi w} \sin \frac{n\pi w}{p} \tag{13}$$

and

$$R_n = \frac{k_z}{k_{zn}} \delta_{n0} + i \frac{k}{k_{zn}} \frac{\tan kd}{n\pi} \sin\left(\frac{n\pi w}{p}\right) G_0 \tag{14}$$

Substituting (14) into (13), we find

$$G_0 = 2 \left[1 - i \sum_{n=-\infty}^{\infty} \frac{kp \tan kd}{(n\pi)^2 k_{zn} w} \sin^2 \frac{n\pi w}{p} \right]^{-1} \tag{15}$$

Substituting back into (14) we obtain the reflection coefficients R_n.

This mode-matching technique is often useful in solving problems involving periodic structures. The use of Floquet modes also greatly facilitates the discussions of scattered waves. As another example, consider a similar structure made of parallel conducting plates of widths $(p-w)$ and separated by distance w. The conducting plane at $z = -d$ in Figure 6.3.1 is now removed. For an incident TM wave with the magnetic field in the $\hat{y}$ direction, the TEM waveguide mode in the parallel-plate regions is excited. The reflectivity is always less than unity. For an incident TE wave with electric field $\overline{E}$ in the $\hat{y}$ direction, the excited guided waves in the plate regions are all TE modes. Thus, if the plate separation is such that $kw < \pi$, all guided-wave modes are evanescent and all the incident power will be scattered.

b. Scattering by Periodic Dielectric Surfaces

Consider a plane wave incident on a periodic surface described by $f(x) = f(x + p)$ where p denotes the period of the surface in the $\hat{x}$ direction [Fig. 6.3.2]. Let the incident field be

$$\overline{E}_i = \hat{y} E_{iy}(\bar{r}) = \hat{y} E_o e^{i\bar{k}_i \cdot \bar{r}} \tag{16}$$

where $\bar{k}_i = \hat{x} k_{ix} - \hat{z} k_{iz}$ is the incident wave vector.

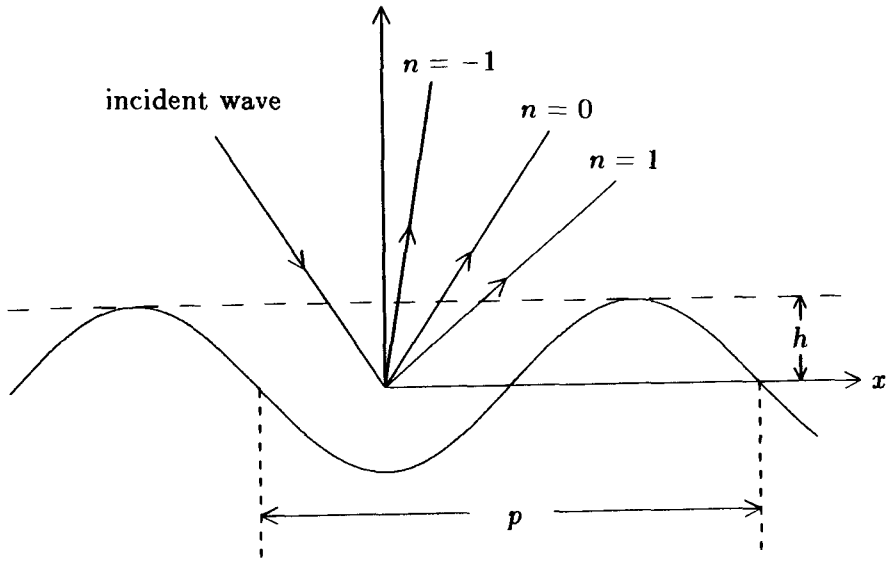


Figure 6.3.2 Scattering by a periodic rough surface.

We make use of the scalar formulation of the Huygens' principle to derive the extinction theorem for the solution of the problem. From equation (5.3.29) and using the scalar Green's function in region 0, we have the total electric field in regions 0 and 1 as follows:

$$E_{iy}(\bar{r}) + \iint_{-\infty}^{\infty} dS' \left\{ E_y(\bar{r}')\hat{n} \cdot \nabla_s' g(\bar{r}, \bar{r}') - g(\bar{r}, \bar{r}')\hat{n} \cdot \nabla_s' E_y(\bar{r}') \right\} = \begin{cases} E_y(\bar{r}) & z > f(x) \\ 0 & z < f(x) \end{cases} \tag{17}$$

where

$$dS'\,\hat{n} = \left[\hat{z} - \hat{x} \frac{d}{dx'} f(x') \right] dx' \tag{18}$$

Equation (17) states that the total field in region 0 is equal to the sum of the incident field and the scattered field produced by the induced surface currents. The sum of the incident field and the scattered field in region 1 is equal to zero on account of Huygens' principle. This is referred to as the *extinction theorem.*

The surface integral in (17) is over an infinite domain. However, it can be condensed into a single period. First we recognize that the surface fields have periodic properties, e.g.,

$$E_y(\bar{r}' + \hat{x}np) = E_y(\bar{r}')e^{ik_{ix}np} \tag{19}$$

From the theory of Fourier series, we can represent a periodic train of Dirac delta functions by an infinite summation of complex exponentials

$$\sum_{m=-\infty}^{\infty} e^{i(k_{ix}-k_x)mp} = \sum_{m=-\infty}^{\infty} \frac{2\pi}{p}\delta\left(k_x - k_{ix} - \frac{2m\pi}{p}\right) \tag{20}$$

The Green's function for region 0 is

$$g(\bar{r},\bar{r}') = \frac{i}{4}H_0^{(1)}(k|\bar{r}-\bar{r}'|) = \frac{i}{4\pi}\int_{-\infty}^{\infty} dk_x \frac{1}{k_z} e^{ik_x(x-x')+ik_z|z-z'|} \tag{21}$$

To reduce the surface integal in (17) to only one period, we use the following identity derived from (20) and (21), with the translation phase factor $e^{ik_{ix}mp}$ in (19) included in the Green's function,

$$\begin{aligned}\sum_{m=-\infty}^{\infty} g(\bar{r},\bar{r}'+\hat{x}mp)e^{ik_{ix}mp} &= \frac{i}{4\pi}\int_{-\infty}^{\infty} dk_x \frac{1}{k_z} e^{ik_x(x-x')+ik_z|z-z'|}\sum_m e^{i(k_{ix}-k_x)mp} \\ &= g_p(\bar{r},\bar{r}') = \frac{i}{2p}\sum_n \frac{1}{k_{zn}} e^{ik_{xn}(x-x')+ik_{zn}|z-z'|}\end{aligned} \tag{22}$$

where

$$k_{xn} = k_{ix} + n\frac{2\pi}{p} \tag{23a}$$

$$k_{zn} = (k^2 - k_{xn}^2)^{1/2} \tag{23b}$$

When k_{xp}^2 is larger than k^2, we choose $k_{zn} = i(k_{xn}^2 - k^2)^{1/2}$. Making use of (19) and (22) in (17) we obtain

$$\begin{aligned}E_{iy}(\bar{r}) + \int_p dS' \left\{E_y(\bar{r}')\hat{n}\cdot\nabla_s' g_p(\bar{r},\bar{r}') - g_p(\bar{r},\bar{r}')\hat{n}\cdot\nabla_s' E_y(\bar{r}')\right\} \\ = \begin{cases} E_y(\bar{r}) & z > f(x) \\ 0 & z < f(x)\end{cases}\end{aligned} \tag{24}$$

This integration now only extends over a single period.

In terms of the Green's function (24) for region 1, we find, similar to (22) and (24),

$$\sum_{m=-\infty}^{\infty} g_1(\bar{r},\bar{r}' + \hat{x}mp)e^{ik_{iz}mp} = g_{1p}(\bar{r},\bar{r}')$$

$$= \frac{i}{2p}\sum_n \frac{1}{k_{1zn}} e^{ik_{xn}(x-x')+ik_{1zn}|z-z'|} \tag{25}$$

where

$$k_{1zn} = (k_1^2 - k_{xn}^2)^{1/2} \tag{26}$$

and

$$-\int_p dS' \left\{E_{1y}(\bar{r}')\hat{n}\cdot\nabla'_s g_{1p}(\bar{r},\bar{r}') - g_{1p}(\bar{r},\bar{r}')\hat{n}\cdot\nabla'_s E_{1y}(\bar{r}')\right\}$$

$$= \begin{cases} 0 & z > f(x) \\ E_{1y}(\bar{r}) & z < f(x) \end{cases} \tag{27}$$

Notice that the minus sign in front of the integral signifies the fact that the unit vector $\hat{n}$ is the surface normal pointing into region 0 and Huygens' principle requires the surface normal to point outward.

From the representations of the periodic Green's functions in (22) and (25), we see that the scattered waves are propagating in discrete Floquet modes, with directions determined by (23) and (26). Defining θ_{rn} as the angle of the n th-order reflected Floquet mode and θ_{tn} as the angle of the n th-order transmitted Floquet mode, we find

$$k\sin\theta_{rn} = k_{xn} = k\sin\theta_i + n\frac{2\pi}{p} \tag{28a}$$

$$k_1\sin\theta_{tn} = k_{txn} = k\sin\theta_i + n\frac{2\pi}{p} \tag{28b}$$

The phase-matching results of (28) are depicted in Figure 6.3.3. Given the incident wave vector $\overline{k_0}$ as shown in the figure, the zeroth order modes of the reflected wave and the transmitted wave are the same as those for a planar surface.

Equation (24) can be simplified by allowing z to be larger than f_{max} or smaller than f_{min} where f_{max} and f_{min} are respectively the maximum and minimum values of the surface profile $f(x)$. For

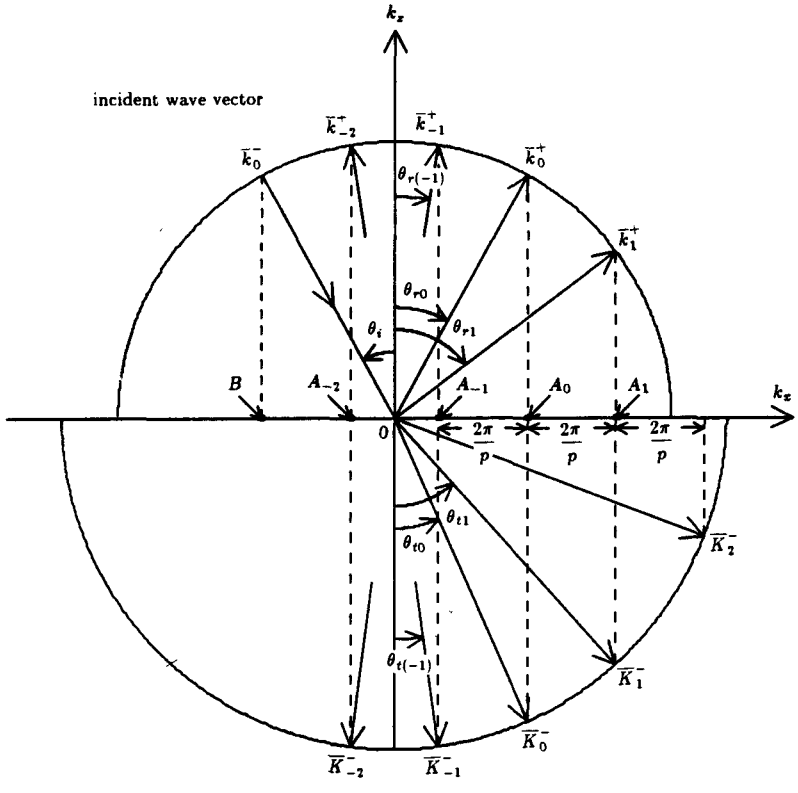


Figure 6.3.3 k **space diagram.**

$z > f_{max}$, $|z-z'|$ becomes $z-z'$ and for $z < f_{min}$, $|z-z'|$ becomes $-(z-z')$. Thus

$$E_y(\bar{r}) = E_{iy}(\bar{r}) + \sum_n b_n \frac{e^{i\bar{k}_n^{+}\cdot\bar{r}}}{\sqrt{k_{zn}}} \qquad z > f_{max} \tag{29a}$$

$$0 = E_{iy}(\bar{r}) - \sum_n a_n \frac{e^{i\bar{k}_n^{-}\cdot\bar{r}}}{\sqrt{k_{zn}}} \qquad z < f_{min} \tag{29b}$$

where

$$\bar{k}_n^{\pm} = \hat{x}k_{xn} \pm \hat{z}k_{zn} \tag{30}$$

are the wave vectors of the Floquet modes. The coefficients b_n and a_n are related to the surface field by the following integrals:

$$b_n = \frac{i}{2p}\int_p dS' \left\{ E_y(\bar{r}')\hat{n}\cdot\nabla'_s \frac{e^{-i\bar{k}_n^+\cdot\bar{r}'}}{\sqrt{k_{zn}}} - \frac{e^{-i\bar{k}_n^+\cdot\bar{r}'}}{\sqrt{k_{zn}}}\hat{n}\cdot\nabla'_s E_y(\bar{r}') \right\} \quad (31)$$

and

$$a_n = -\frac{i}{2p}\int_p dS' \left\{ E_y(\bar{r}')\hat{n}\cdot\nabla'_s \frac{e^{-i\bar{k}_n^-\cdot\bar{r}'}}{\sqrt{k_{zn}}} - \frac{e^{-i\bar{k}_n^-\cdot\bar{r}'}}{\sqrt{k_{zn}}}\hat{n}\cdot\nabla'_s E_y(\bar{r}') \right\} \quad (32)$$

Similarly, making use of (27) we find

$$0 = -\sum_n B_n \frac{e^{i\bar{k}_{1n}^+\cdot\bar{r}}}{\sqrt{k_{1zn}}} \qquad z > f_{max} \quad (33a)$$

$$E_{1y}(\bar{r}) = \sum_n A_n \frac{e^{i\bar{k}_{1n}^-\cdot\bar{r}}}{\sqrt{k_{1zn}}} \qquad z < f_{min} \quad (33b)$$

where

$$\bar{k}_{1n}^{\pm} = \hat{x}k_{1xn} \pm \hat{z}k_{1zn} \quad (34)$$

$$B_n = \frac{i}{2p}\int_p dS' \left\{ E_{1y}(\bar{r}')\hat{n}\cdot\nabla'_s \frac{e^{-i\bar{k}_{1n}^+\cdot\bar{r}'}}{\sqrt{k_{1zn}}} - \frac{e^{-i\bar{k}_{1n}^+\cdot\bar{r}'}}{\sqrt{k_{1zn}}}\hat{n}\cdot\nabla'_s E_{1y}(\bar{r}') \right\} \quad (35)$$

$$A_n = -\frac{i}{2p}\int_p dS' \left\{ E_{1y}(\bar{r}')\hat{n}\cdot\nabla'_s \frac{e^{-i\bar{k}_{1n}^-\cdot\bar{r}'}}{\sqrt{k_{1zn}}} - \frac{e^{-i\bar{k}_{1n}^-\cdot\bar{r}'}}{\sqrt{k_{1zn}}}\hat{n}\cdot\nabla'_s E_{1y}(\bar{r}') \right\} \quad (36)$$

Equations (29*b*) and (33*a*) are referred to as the *extended boundary conditions* (EBC). By letting the observation point be outside the trough regions of the periodic surface, we find from (29*b*)

$$E_{iy}(\bar{r}) = \sum_n a_n \frac{e^{i\bar{k}_n^-\cdot\bar{r}}}{\sqrt{k_{zn}}} \qquad z < f_{min}$$

which gives

$$a_n = \delta_{no}\sqrt{k_z}E_o \quad (37)$$

and from (33*a*)

$$B_n = 0 \quad (38)$$

Knowing a_n and B_n, we can solve (32) and (35) for the unknown surface fields which can then be used to determine the scattered field amplitudes b_n and A_n from (31) and (36).

We now apply the boundary conditions for the tangential fields on the periodic surface S'. The continuity of tangential electric fields requires

$$E_y(x, z = f(x)) = E_{1y}(x, z = f(x)) \tag{39}$$

and the continuity of tangential magnetic fields requires

$$\hat{n} \times \nabla_s \times \hat{y} E_y = \hat{n} \times \nabla_s \times \hat{y} E_{1y}$$

or equivalently

$$\hat{n} \cdot \nabla_s E_y = \hat{n} \cdot \nabla_s E_{1y} \tag{40}$$

where

$$\hat{n} = \frac{\hat{z} - \hat{x}\dfrac{df}{dx}}{\sqrt{1 + (df/dx)^2}}$$

$$\hat{n} dS = dx \left[\hat{z} - \hat{x}\frac{df(x)}{dx}\right]$$

Noting that the surface fields are dependent on x only, we express the unknown fields in terms of Fourier series expansions as follows:

$$E_y(\bar{r}) = \sum_n 2\alpha_n^s e^{i(k_{ix} + nK)x} \tag{41}$$

$$dS\, \hat{n} \cdot \nabla_s E_y(\bar{r}) = -i dx \sum 2\beta_n^s e^{i(k_{ix} + nK)x} \tag{42}$$

where $K = 2\pi/p$. This choice of basis functions is appropriate because the surface fields when multiplied by the term $\exp(-ik_{ix}x)$ are periodic functions of x [Chuang and Kong, 1983].

Substituting (41) and (42) into (32) and defining $Q_{D1}^{\pm}$ and $Q_{N1}^{\pm}$ as the Dirichlet and Neumann matrices with elements

$$\left[\overline{\overline{Q}}_D^{\pm}\right]_{mn} = \pm\frac{1}{p}\int_p dx \frac{e^{-ik_m^{\pm}\cdot\bar{r}}}{\sqrt{k_{zm}}} e^{i[k_{ix} + nK]x}$$

$$= \pm\frac{1}{p\sqrt{k_{zm}}}\int_p dx\, e^{-i[(m-n)Kx \pm k_{zm} f(x)]} \tag{43a}$$

$$\left[\overline{\overline{Q}}_N^{\pm}\right]_{mn} = \pm\frac{(-i)}{p}\int_p dS\,\hat{n}\cdot\left(\nabla_s \frac{e^{-ik_m^{\pm}\cdot\bar{r}}}{\sqrt{k_{zm}}}\right) e^{i[k_{ix}+nK]x}$$

$$= \pm\frac{(-1)}{p\sqrt{k_{zm}}}\int_p dx\left[k_{xm}\frac{df}{dx} \mp k_{zm}\right] e^{-i[(m-n)Kx \pm k_{zm}f(x)]}$$

$$= \pm\frac{(-1)}{p\sqrt{k_{zm}}}\int_p dx\left[\frac{k_{xm}(m-n)K}{k_{zm}} \mp k_{zm}\right] e^{-i[(m-n)Kx \pm k_{zm}f(x)]}$$

$$= \pm\frac{(-k^2 + k_{xm}k_{xn})}{k_{zm}p\sqrt{k_{zm}}}\int_p dx\, e^{-i[(m-n)Kx \pm k_{zm}f(x)]} \tag{43b}$$

we obtain the matrix equation

$$\bar{a} = -\overline{\overline{Q}}_D^{-}\cdot\bar{\beta}^s - \overline{\overline{Q}}_N^{-}\cdot\bar{\alpha}^s \tag{44}$$

Similarly, (35) leads to the matrix equation

$$-\overline{\overline{Q}}_{D1}^{+}\cdot\bar{\beta}^s - \overline{\overline{Q}}_{N1}^{+}\cdot\bar{\alpha}^s = \bar{B} = 0 \tag{45}$$

where

$$\left[Q_{D1}^{\pm}\right]_{mn} = \pm\frac{1}{p\sqrt{k_{1zm}}}\int_p dx\, e^{-i[(m-n)Kx \pm k_{1zm}f(x)]} \tag{46a}$$

$$\left[Q_{N1}^{\pm}\right]_{mn} = \pm\frac{(-k_1^2 + k_{1xm}k_{1xn})}{k_{1zm}\sqrt{k_{1zm}}\,p}\int_p dx\, e^{-i[(m-n)Kx \pm k_{1zm}f(x)]} \tag{46b}$$

Combining (45) and (46), we can determine $\bar{\alpha}^s$ and $\bar{\beta}^s$ from the following matrix equation

$$-\begin{bmatrix} \overline{\overline{Q}}_N^{-} & \overline{\overline{Q}}_D^{-} \\ \overline{\overline{Q}}_{N1}^{+} & \overline{\overline{Q}}_{D1}^{+} \end{bmatrix}\begin{bmatrix} \bar{\alpha}^s \\ \bar{\beta}^s \end{bmatrix} = \begin{bmatrix} \bar{a} \\ 0 \end{bmatrix} \tag{47}$$

Notice that $\bar{a}$ on the right-hand side has been calculated and is given by (37).

After $\bar{\alpha}^s$ and $\bar{\beta}^s$ are determined from (47), the scattered field amplitudes can be obtained from (31) and (36) with the Q matrices defined in (43) and (46). The upward-propagating field amplitudes are

$$\bar{b} = -\overline{\overline{Q}}_D^{+}\cdot\bar{\beta}^s - \overline{\overline{Q}}_N^{+}\cdot\bar{\alpha}^s \tag{48}$$

The downward-propagating field amplitudes are

$$\overline{A} = -\overline{\overline{Q}}_{D1}^{-} \cdot \overline{\beta}^{s} - \overline{\overline{Q}}_{N1}^{-} \cdot \overline{\alpha}^{s} \tag{49}$$

Thus the problem of scattering of the scattering of a TE wave incident upon a periodic dielectric medium is now completely solved. The solution for a TM incident wave can be carried out in a similar manner.

For a sinusoidal rough surface with

$$f(x) = -h\cos\left(\frac{2\pi}{p}x\right) \tag{50}$$

the $\overline{\overline{Q}}^{\pm}$ matrices can be calculated by carrying out the integrations in (43) and (47). Expressed in terms of Bessel functions, we have

$$\left[\overline{\overline{Q}}_{D}^{\pm}\right]_{mn} = \pm\frac{1}{\sqrt{k_{zm}}}(\pm i)^{|m-n|}\, J_{|m-n|}(k_{zm}h) \tag{51a}$$

$$\left[\overline{\overline{Q}}_{N}^{\pm}\right]_{mn} = \left(\frac{-k^2 + k_{xm}k_{xn}}{k_{zm}\sqrt{k_{zm}}}\right)(\pm i)^{|m-n|}\, J_{|m-n|}(k_{zm}h) \tag{51b}$$

$$\left[\overline{\overline{Q}}_{D1}^{\pm}\right]_{mn} = \pm\frac{1}{\sqrt{k_{1zm}}}(\pm i)^{|m-n|}\, J_{|m-n|}(k_{1zm}h) \tag{51c}$$

$$\left[\overline{\overline{Q}}_{N1}^{\pm}\right]_{mn} = \frac{(-k_1^2 + k_{xm}k_{xn})}{k_{1zm}\sqrt{k_{1zm}}}(\pm i)^{|m-n|}\, J_{|m-n|}(k_{1zm}h) \tag{51d}$$

For a general profile of a periodic surface defined by a single-valued function $z = f(x)$, the EBC method can be applied by numerically integrating (43) and (46) to calculate the elements of the $\overline{\overline{Q}}^{\pm}$ matrices. In practice, the matrices used may become ill-conditioned when the surface corrugation is deep or when the corrugation depth divided by the period is large.

6.4 Scattering by Periodic Media

In the study of periodic media with applications to holography, to ultrasonic light diffraction, and to various active and passive components in integrated optics, the coupled-mode approach proves to be the simplest method leading to results that can be easily interpreted

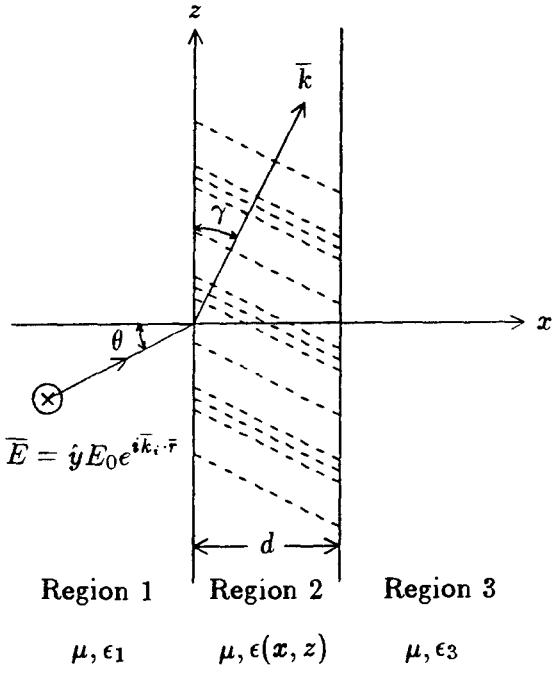


Figure 6.4.1 Geometrical configuration of the problem.

physically. Consider a periodic medium (Figure 6.4.1 without regions 1 and 3) described by the permittivity

$$\epsilon(x, z) = \epsilon_2(1 + \eta \cos(\overline{K} \cdot \overline{r})) \tag{1}$$

where $\overline{K} = K(\hat{x} \sin\gamma + \hat{z} \cos\gamma)$, $K = 2\pi/\Lambda$, and Λ is the periodicity. For a TE wave polarized in the $\hat{y}$ direction, the electric field vector $\overline{E}$ satisfies the wave equation

$$\left[\nabla^2 + \omega^2 \mu \epsilon(x, z)\right] E_y(x, z) = 0 \tag{2}$$

To facilitate the derivation of the coupled-mode equations, we let

$$E_y(x, z) = \sum_{m=-\infty}^{\infty} \phi_m(x) e^{im\pi/2} e^{ik_{mx}x} \tag{3}$$

where

$$k_{mx} = k_{0x} + mK \cos\gamma \tag{4}$$

k_{0x} is the x component of the wave vector for the zeroth-order Floquet mode.

Substituting (3) in (2) and letting $\gamma = 0$ yield

$$\frac{d^2\phi_m}{dx^2} + \left(k_2^2 - k_{mx}^2\right)\phi_m + i\frac{\eta}{2}k_2^2\left(\phi_{m+1} - \phi_{m-1}\right) = 0 \tag{5}$$

where $k_2 = \omega(\mu\epsilon)^{1/2}$. Equation (5) represents a set of coupled second-order differential equations.

Near the first Bragg angle, the two Floquet modes that couple strongly to each other are of the zeroth order and first order. Keeping only these two modes, (5) becomes

$$\frac{d^2\phi_0}{dx^2} + \left(k_2^2 - k_{0x}^2\right)\phi_0 = i\frac{\eta k_2^2}{2}\phi_{-1} \tag{6a}$$

$$\frac{d^2\phi_{-1}}{dx^2} + \left(k_2^2 - k_{-1x}^2\right)\phi_{-1} = -i\frac{\eta k_2^2}{2}\phi_0 \tag{6b}$$

This set of coupled equations can be converted into two uncoupled Helmholtz equations. We let

$$U_1 = \phi_0 + i\alpha_1\phi_{-1} \tag{7a}$$

$$U_2 = \phi_0 + i\alpha_2\phi_{-1} \tag{7a}$$

or, equivalently,

$$\phi_0 = \frac{\alpha_2 U_1 - \alpha_1 U_2}{\alpha_2 - \alpha_1} \tag{7c}$$

$$\phi_{-1} = i\frac{U_1 - U_2}{\alpha_2 - \alpha_1} \tag{7d}$$

where α_1 and α_2 are constants to be determined. Adding (6a) to (6b) multiplied by α_j, where $j = 1, 2$, we obtain

$$\frac{d^2U_j}{dx^2} + \left(k_2^2 - k_{0x}^2 - \alpha_j\frac{\eta k_2^2}{2}\right)U_j = 0 \tag{8}$$

and

$$\alpha_j = (1/\eta k_2^2)\left\{k_{-1x}^2 - k_{0x}^2 \pm \left[(k_{-1x}^2 - k_{0x}^2)^2 + \eta^2 k_2^4\right]^{1/2}\right\} \tag{9}$$

The plus sign in front of the radical is for α_1 and the minus sign for α_2. Clearly, exponential functions are solutions to (8). The four

independent solutions to the two coupled-mode equations in (6) are therefore

$$\phi_0 = \frac{1}{\alpha_2 - \alpha_1}\left[\alpha_2\begin{pmatrix} W\exp(ik_{2x}^a x) \\ X\exp(-ik_{2x}^a x)\end{pmatrix} - \alpha_1\begin{pmatrix} Y\exp(ik_{2x}^b x) \\ Z\exp(-ik_{2x}^b x)\end{pmatrix}\right] \tag{10a}$$

$$\phi_{-1} = \frac{i}{\alpha_2 - \alpha_1}\left[\begin{pmatrix} W\exp(ik_{2x}^a x) \\ X\exp(-ik_{2x}^a x)\end{pmatrix} - \begin{pmatrix} Y\exp(ik_{2x}^b x) \\ Z\exp(-ik_{2x}^b x)\end{pmatrix}\right] \tag{10b}$$

where

$$k_{2x}^a = [(1 - \alpha_1\eta/2)k_2^2 - k_{0x}^2]^{1/2} \tag{11a}$$

$$k_{2x}^b = [(1 - \alpha_2\eta/2)k_2^2 - k_{0x}^2]^{1/2} \tag{11b}$$

The constants W, X, Y, and Z are to be determined by the appropriate boundary conditions for specific problems.

a. First-Order Coupled-Mode Equations

First-order coupled-mode equations are easily obtained from (2) and (3) by letting

$$\phi_m(x) = \psi_m(x)e^{ik_{mx}x} \tag{12}$$

where

$$k_{mx} = k_{0x} + mK\sin\gamma \tag{13}$$

and

$$k_{0x}^2 + k_{0z}^2 = k_2^2 \tag{14}$$

Neglecting the second-order derivative term, we obtain

$$\frac{d\psi_m(x)}{dx} + i\frac{mK}{2k_{mx}}[mK + 2(k_{0x}\cos\gamma + k_{0x}\sin\gamma)] + \frac{\eta k_2^2}{4k_{mx}}[\psi_{m+1}(x) - \psi_{m-1}(x)] = 0 \tag{15}$$

When $\gamma = 0$, it becomes the differential equation used by Klein and Cook [1967] in their numerical solution by converting (15) to difference equations.

The Raman-Nath regime is characterized by $\gamma = 0$, $k_{0x} = 0$, $k_{mx} \approx k_2$, and neglecting the ψ_m term in (15). These conditions are

met when a wave is normally incident on a periodic medium with periodicity along the $\hat{z}$ direction and with a very small K. Equation (15) becomes

$$\frac{d\psi_m(x)}{dx} + \frac{\eta k_2}{4}[\psi_{m+1}(x) - \psi_{m-1}(x)] = 0 \tag{16}$$

This equation is identical to the recurrence relation for Bessel functions. Thus

$$\psi_m(x) = J_m(\eta k_2 x/2) \tag{17}$$

Note that $J_0^2(x) + 2\sum_{m=1}^{\infty} J_m^2(x) = 1$, which is a statement of conservation of energy.

We now assume that only two modes exist, $m = 0$ and $m = -1$. Then (15) becomes

$$\frac{d\psi_0}{dx} - \frac{\eta k_2^2}{4k_{0x}}\psi_{-1} = 0 \tag{18a}$$

$$\frac{d\psi_{-1}}{dx} + \frac{K}{4k_{-1x}}[K - 2(k_{0x}\cos\gamma + k_{0x}\sin\gamma)]\psi_{-1}(x) + \frac{\eta k_2^2}{4k_{-1x}}\psi_0 = 0 \tag{18b}$$

In Kogelnik's [1969] treatment of thick holograms, he requires that for transmission holograms $\phi_0(x = 0) = 1$, and $\phi_{-1}(x = 0) = 0$, whereas for reflection holograms $\phi_0(x = 0) = 1$ and $\phi_{-1}(x = d) = 0$, where d is the thickness of the hologram. Only two boundary conditions are required because the two governing equations involve only first-order derivatives.

The Phariseau limit is characterized by $\gamma = 0$ and $k_{0x} = K/2$. This condition occurs when a wave propagates at the Bragg angle in a periodic medium with periodicity along the $\hat{z}$ direction. Noting that from (4), $k_{-1x} = k_{0x}$, we obtain from (18)

$$k_{0x}\frac{d\psi_0(x)}{dx} = \frac{\eta k_2^2}{4}\psi_{-1}(x) \tag{19a}$$

$$k_{0x}\frac{d\psi_{-1}(x)}{dx} = -\frac{\eta k_2^2}{4}\psi_0(x) \tag{19b}$$

Imposing the boundary condition $\psi_0(x = 0) = 1$ and $\psi_{-1}(x = 0) = 0$, we have the solution

$$\psi_0(x) = \cos(\eta k_2^2 x/4k_{0x}) \tag{20a}$$

$$\psi_{-1}(x) = \sin(\eta k_2^2 x/4k_{0x}) \tag{20b}$$

Conservation of energy holds for all x since $\psi_0^2(x) + \psi_{-1}^2(x) = 1$.

In order to reduce to the first-order coupled-mode equations used by Kogelnik and Shank [1971] in their treatment of distributed feedback devices, we let $\gamma = \pi/2$, $k_{mx} = 0$, and $k_{0x} = k_2$. This occurs when the spatial periodicity is along the $\hat{x}$ direction and the wave is also propagating along the $\hat{x}$ direction. We let

$$E(x,z) = \Psi_0(x)e^{i(K/2)x} + \Psi_{-1}(x)e^{-i(K/2)x} \tag{21}$$

Substituting in (2), and neglecting the second-order derivative term, we have

$$\Psi_0'(x) - (2i/K)(k_2^2 - K^2/4)\Psi_0(x) + (\eta k_2^2/K)\Psi_{-1}(x) = 0 \tag{22a}$$
$$-\Psi_{-1}'(x) - (2i/K)(k_2^2 - K^2/4)\Psi_{-1}(x) + (\eta k_2^2/K)\Psi_0(x) = 0 \tag{22b}$$

This set of coupled-mode equations has been extensively studied by Kogelnik and Shank. The wavenumber $k_2 = \omega(\mu\epsilon)^{1/2}$ is made a variable that corresponds to a changing angular frequency ω. Right at the first Bragg frequency, $k_2 = K/2$ and (22) becomes the coupled equations shown in (19).

b. Reflection and Transmission by Periodically-Modulated Slab

Consider a periodic slab medium with thickness d [Fig. 6.4.1]. A plane wave is incident at the slab with wave vector $\bar{k} = \hat{x}k_x + \hat{z}k_{0z}$, where $k_{0z} = k_1 \sin\theta$, $k_1 = \omega(\mu\epsilon_1)^{1/2}$, and θ is the angle of incidence. The electric field for the reflected wave takes the form

$$E_y = E_0 \exp(ik_{1x}^a x + ik_{0z}z) + R_0 \exp(-ik_{1x}^a x + ik_{0z}z) + R_{-1} \exp(-k_{1x}^b x + ik_{-1z}z) \tag{23}$$

where

$$k_{1x}^a = (k_1^2 - k_{0z}^2)^{1/2} \tag{24a}$$
$$k_{1x}^b = (k_1^2 - k_{-1z}^2)^{1/2} \tag{24b}$$

and R_0 and R_1 are the reflection coefficients for the zeroth- and the first-order modes. The transmitted wave takes the form

$$E_y = T_0 \exp[ik_{3x}^a + ik_{0z}z] + T_{-1} \exp[ik_{3x}^b x + ik_{-1z}z] \tag{25}$$

where

$$k_{3x}^a = (k_3^2 - k_{0x}^2)^{1/2} \tag{26a}$$

$$k_{3x}^b = (k_3^2 - k_{-1x}^2)^{1/2} \tag{26b}$$

and T_0 and T_{-1} are the transmission coefficients for the zeroth- and the first-order modes.

Inside the slab medium, the electric field E_y takes the form

$$\begin{aligned} E_y = {} & \phi_0 e^{ik_{0x}z} + i\phi_{-1} e^{ik_{-1}z} = [1/(\alpha_2 - \alpha_1)] \\ & \cdot (\alpha_2 W e^{ik_{2x}^a x} + \alpha_2 X e^{-ik_{2x}^a x} - \alpha_1 Y e^{ik_{2x}^b x} - \alpha_1 Z e^{-ik_{2x}^b x}) e^{ik_{0x}z} \\ & - [1/(\alpha_2 - \alpha_1)](W e^{ik_{2x}^a x} + X e^{-ik_{2x}^a x} - Y e^{ik_{2x}^b x} - Z e^{-ik_{2x}^b x}) \\ & \cdot e^{ik_{-1x}z} \end{aligned} \tag{27}$$

Thus we have a total of eight constants R_0, R_{-1}, T_0, T_{-1}, W, X, Y, and Z to be determined by the boundary conditions, which require that the tangential electric and magnetic fields be continuous at the boundaries $x = 0$ and $x = d$. The tangential magnetic field H_z is determined from E_y by using the Maxwell equation

$$H_z = \frac{1}{i\omega\mu} \frac{\partial}{\partial x} E_y \tag{28}$$

The four boundary conditions must be met for all z. This results in eight linear simultaneous equations to be solved for the unknown coefficients.

After considerable algebra, we obtain the transmission and reflection coefficients as follows:

$$T_0 = \frac{4(\alpha_2 - \alpha_1)k_{1x}^a(\alpha_1 A_{bb} - \alpha_2 B_{bb})e^{-ik_{3x}^a d}}{(\alpha_2 A_{aa} - \alpha_1 B_{aa})(\alpha_1 A_{bb} - \alpha_2 B_{bb}) - \alpha_1\alpha_2(A_{ab} - B_{ab})(A_{ba} - B_{ba})} \tag{29}$$

$$T_{-1} = \frac{4(\alpha_2 - \alpha_1)k_{1x}^a(A_{ba} - B_{ba})e^{-ik_{3x}^b d}}{(\alpha_2 A_{aa} - \alpha_1 B_{aa})(\alpha_1 A_{bb} - \alpha_2 B_{bb}) - \alpha_1\alpha_2(A_{ab} - B_{ab})(A_{ba} - B_{ba})} \tag{30}$$

$$R_0 = \frac{\alpha_1\alpha_2(\alpha_{ab} - \beta_{ab})(A_{ba} - B_{ba}) - (\alpha_2\alpha_{aa} - \alpha_1\beta_{aa})(\alpha_1 A_{bb} - \alpha_2 B_{bb})}{(\alpha_2 A_{aa} - \alpha_1 B_{aa})(\alpha_1 A_{bb} - \alpha_2 B_{bb}) - \alpha_1\alpha_2(A_{ab} - B_{ab})(A_{ba} - B_{ba})} \tag{31}$$

$$R_{-1} = \frac{(\alpha_2 A_{aa} - \alpha_1 B_{aa})(\alpha_{ab} - \beta_{ab}) - (A_{ab} - B_{ab})(\alpha_2\alpha_{aa} - \alpha_1\beta_{aa})}{(\alpha_2 A_{aa} - \alpha_1 B_{aa})(\alpha_1 A_{bb} - \alpha_2 B_{bb}) - \alpha_1\alpha_2(A_{ab} - B_{ab})(A_{ba} - B_{ba})} \tag{32}$$

where

$$A_{\rho\sigma} = k_{2x}^a\left(1 + \frac{k_{1x}^\rho}{k_{2x}^a}\right)\left(1 + \frac{k_{3x}^\rho}{k_{2x}^a}\right)(e^{-ik_{2x}^a d} - R_{21}^{a\rho} R_{23}^{a\sigma} e^{ik_{2x}^a d}) \tag{33}$$

$$B_{\rho\sigma} = k_{2x}^b\left(1 + \frac{k_{1x}^\rho}{k_{2x}^b}\right)\left(1 + \frac{k_{3x}^\rho}{k_{2x}^b}\right)(e^{-ik_{2x}^b d} - R_{21}^{b\rho} R_{23}^{b\sigma} e^{ik_{2x}^b d}) \tag{34}$$

$$\alpha_{\rho\sigma} = k_{2x}^a\left(1 + \frac{k_{1x}^\rho}{k_{2x}^a}\right)\left(1 + \frac{k_{3x}^\rho}{k_{2x}^a}\right)(R_{21}^{a\rho} e^{-ik_{2x}^a d} - R_{23}^{a\sigma} e^{ik_{2x}^a d}) \tag{35}$$

$$\beta_{\rho\sigma} = k_{2x}^b\left(1 + \frac{k_{1x}^\rho}{k_{2x}^b}\right)\left(1 + \frac{k_{3x}^\rho}{k_{2x}^b}\right)(R_{21}^{b\rho} e^{-ik_{2x}^b d} - R_{23}^{b\sigma} e^{ik_{2x}^b d}) \tag{36}$$

$$R_{ij}^{\rho\alpha} = \frac{k_{ix}^\rho - k_{jx}^\sigma}{k_{ix}^\rho + k_{jx}^\sigma} \tag{37}$$

The superscripts ρ and σ stand for either a or b and the subscripts i and j for 1, 2, or 3.

The solutions (29)–(32) reduce to known results in the absence of modulation, where $A_{\rho\sigma} = B_{\rho\sigma}$, $\alpha_{\rho\sigma} = \beta_{\rho\sigma}$, and

$$R_{ij}^{\rho\sigma} = R_{ij} = \frac{1 - k_{jx}/k_{ix}}{1 + k_{jx}/k_{ix}}$$

we have $R_{-1} = T_{-1} = 0$ and

$$R_0 = \frac{-R_{21} + R_{23}e^{i2k_{2x}d}}{1 - R_{21}R_{23}e^{i2k_{2x}^a d}} \tag{38}$$

$$T_0 = \frac{4\exp[i(k_{2x} - k_{3x})d]}{(1 + k_{2x}/k_{1x})(1 + k_{3x}/k_{2x})(1 - R_{21}R_{23}e^{i2k_{2x}^a d})} \tag{39}$$

It is easily shown that $|R_0|^2 + k_{3x}|T_0|^2/k_{1x} = 1$, which is the statement for power conservation.

When modulation is present but the wave is incident at exactly the Bragg angle, $K_{0z} = K/2$ and $\alpha_1 = -\alpha_2 = 1$. The expressions for the transmission and reflection coefficients can be simplified to read

$$R_0 = \frac{1}{2}(R_a + R_b) \tag{40}$$

$$R_{-1} = \frac{1}{2}(R_a - R_b) \tag{41}$$

$$T_0 = \frac{1}{2}(T_a + T_b) \tag{42}$$

$$T_{-1} = \frac{1}{2}(T_a + T_b) \tag{43}$$

where

$$R_a = \frac{-R_{21}^a + R_{23}^a e^{i2k_{2z}^a d}}{1 - R_{21}^a R_{23}^a e^{i2k_{2z}^a d}} \tag{44}$$

$$R_b = \frac{-R_{21}^b + R_{23}^b e^{i2k_{2z}^b d}}{1 - R_{21}^b R_{23}^b e^{i2k_{2z}^b d}} \tag{45}$$

$$T_a = \frac{4\exp\left[i(k_{2x}^a - k_{3x})d\right]}{(1 + k_{2x}^a/k_{1x})(1 + k_{3x}/k_{2x}^a)(1 - R_{21}^a R_{23}^a e^{i2k_{2z}^a d})} \tag{46}$$

$$T_b = \frac{4\exp\left[i(k_{2x}^b - k_{3x})d\right]}{(1 + k_{2x}^b/k_{1x})(1 + k_{3x}/k_{2x}^b)(1 - R_{21}^b R_{23}^b e^{i2k_{2z}^b d})} \tag{47}$$

$$R_{2j}^\sigma = \frac{k_{2x}^\sigma - k_{jx}}{k_{2x}^\sigma + k_{jx}} \tag{48}$$

$$k_{2x}^a = \left[(1 - \frac{1}{2}\eta)k_2^2 - \frac{1}{4}K^2\right]^{1/2} \tag{49}$$

$$k_{2x}^b = \left[(1 + \frac{1}{2}\eta)k_2^2 - \frac{1}{4}K^2\right]^{1/2} \tag{50}$$

We see that the zeroth-order reflection and transmission coefficients are composed of two terms similar to (38) and (39); one term corresponds to the result of reflection and transmission by a slab with equivalent permittivity $(1 - \eta/2)^{1/2}\epsilon_2$ and the other term to a slab with permittivity $(1 + \eta/2)^{1/2}\epsilon_2$, as seen from (49) and (50).

Notice that the solutions (44)–(47) satisfy conservation of energy, since

$$\begin{aligned} &|R_0|^2 + |R_{-1}|^2 + (k_{3z}/k_{1z})(|T_0|^2 + |T_{-1}|^2) \\ &= \frac{1}{2}\left[|R_a|^2 + (k_{3z}/k_{1z})|T_a|^2 + |R_b|^2 + (k_{3z}/k_{1z})|T_b|^2\right] = 1 \end{aligned} \tag{51}$$

This is also observed from (3), (6), and (28) by requiring that the spatial derivative of the time-average Poynting power density in the $\hat{x}$ direction be zero.

c. *Far-Field Diffraction of a Gaussian Beam*

The electric field intensity of a transmitted Gaussian beam can be determined from Huygens' principle by using the two-dimensional Green's function

$$\overline{E}_t = \hat{y}\frac{\omega\mu}{2\eta}\int_{-\infty}^{\infty} dz' E_{ap}(z') H_0^{(1)}\{k_3[(z-z')^2+(x-d)^2]\} \tag{52}$$

where η is $(\mu/\epsilon_2)^{1/2}$, $E_{ap}(z')$ is the aperture field at $x = d$ representing either E_0 or E_{-1}, $k_3 = \omega(\mu\epsilon_3)^{1/2}$, and $H_0^{(1)}$ is the zeroth-order Hankel function of the first kind. In the radiation zone, the Fraunhofer approximation leads to

$$\overline{E} = \hat{y}\left[\frac{k_3}{i2\pi(x-d)}\right]^{1/2} e^{ik_3 x + ik_3 z^2/2x}\int_{-\infty}^{\infty} dz' E_{ap}(z') e^{ik_3 z' \sin\theta}$$

which is essentially the Fourier transform of the aperture field.

We shall calculate the far-field pattern of the zeroth-order beam as

$$P_0(\theta) = \int_{-z_m}^{z_m} dz\, E_0(d,z) e^{-ik_3 z\sin\theta} \tag{53}$$

and the far-field pattern of the Bragg-scattered beam as

$$P_{-1}(\theta) = \int_{-z_m}^{z_m} dz\, E_{-1}(d,z) e^{-ik_3 z\sin\theta} \tag{54}$$

where θ is the angle of observation measured from the x axis in the third medium. The fields E_0 and E_{-1} are

$$E_0(x,z) = \int_{-\infty}^{\infty} dk_{0z}\, G(k_{0z}) T_0(k_{0z}) \exp[ik_{3x}^a(x-d) + ik_{0z}z] \qquad x \geq d$$

$$E_{-1}(x,z) = \int_{-\infty}^{\infty} dk_{0z}\, G(k_{0z}) T_{-1}(k_{0z}) \exp[ik_{3x}^b(x-d) - ik_{-1z}z] \qquad x \geq d$$

In (53) and (54), $z_m = d\tan\theta_B + w_0$, where w_0 is the beam width projected along the z axis. This limit is taken because $E_0(d,z;t)$ and $E_{-1}(d,z;t)$ are confined to the region $|z| < d\tan\theta_B + w_0 = z_m$ and the fields are negligibly small for $|z| > z_m$. We have

$$\begin{aligned} P_0(\theta) &= \int_{-z_m}^{z_m} dz \int_{-\infty}^{\infty} dk_{0z}\, G(k_{0z}) T_0(k_{0z}) e^{ik_{0z}z}\, e^{ik_3 z\sin\theta} \\ &= \int_{-\infty}^{\infty} G(k_{0z})\, dk_{0z}\, T_0(k_{0z}) 2z_m \frac{\sin(k_{0z} - k_3\sin\theta)z_m}{(k_{0z} - k_3\sin\theta)z_m} \end{aligned} \tag{55}$$

$$\begin{aligned} P_{-1}(\theta) &= \int_{-z_m}^{z_m} dz \int_{-\infty}^{\infty} dk_{0z}\, G(k_{0z}) T_{-1}(k_{0z}) e^{ik_{-1z}z}\, e^{ik_3 z\sin\theta} \\ &= \int_{-\infty}^{\infty} dk_{0z} G(k_{0z}) T_{-1}(k_{0z}) 2z_m \frac{\sin\{(k_3\sin\theta - [k_{0z} - K]\}\, z_m}{\{k_3\sin\theta - [k_{0z} - K]\}\, z_m} \end{aligned} \tag{56}$$

The factor

$$2z_m \frac{\sin(k_{0z} - k_3\sin\theta)z_m}{(k_{0z} - k_3\sin\theta)z_m}$$

effectively confines the far-field pattern $P_0(\theta)$ into a small angular range centered at the first Bragg angle $\theta = \theta_B = \sin^{-1}(\lambda/2\Lambda\sqrt{\epsilon_3})$. Similarly, the factor

$$2z_m \frac{\sin\{k_3\sin\theta - [k_{0z} - K]\}\, z_m}{\{k_3\sin\theta - [k_{0z} - K]\}\, z_m}$$

effectively confines the far-field pattern $P_{-1}(\theta)$ into a small angular range centered at the negative first Bragg angle $\theta = -\theta_B$. It should be noted that both factors reduce to Dirac delta functions as $z_m \to \infty$ and the results in (55) and (56) become simply $2\pi G T_0$ and $2\pi G T_{-1}$.

6.5 Scattering by Random Media

In the remote sensing of earth terrain media such as snow, ice, and vegetation canopy, the model of a layered medium is often used. In order to account for the scattering from such layered media, its

volume-scattering effects are characterized by a permittivity with a randomly fluctuating part. We write

$$\epsilon_1(\bar{r}) = \epsilon_1 + \epsilon_{1f}(\bar{r}) \tag{1}$$

where ϵ_{1f} is the randomly fluctuating part of the permittivity and ϵ_1 is the mean permittivity such that the ensemble average $< \epsilon_1(\bar{r}) > = \epsilon_1$. In region 1, the equation governing $\overline{E}_1$ can be written as

$$\nabla \times \nabla \times \overline{E}_1(\bar{r}) - k_1^2 \overline{E}_1(\bar{r}) = Q(\bar{r}) \overline{E}_1(\bar{r}) \tag{2}$$

where $k_1^2 = \omega^2 \mu \epsilon_1$ and $Q(\bar{r}) = \omega^2 \mu \epsilon_{1f}$. Thus the random variable $Q(\bar{r}) \overline{E}_1(\bar{r})$ serves as the distributed volume source.

From (2), the electric field $\overline{E}_1$ in region 1 can be solved in terms of the dyadic Green's function $\overline{\overline{G}}_{11}^{(0)}(\bar{r}, \bar{r}')$

$$\overline{E}_1 = \overline{E}_1^{(0)} + \iiint_{V_1} d^3\bar{r}_1 \, \overline{\overline{G}}_{11}^{(0)}(\bar{r}, \bar{r}_1) \cdot Q(\bar{r}_1) \overline{E}_1(\bar{r}_1) \tag{3}$$

where V_1 is the volume for region 1 containing ϵ_{1f}, $\overline{\overline{G}}_{11}^{(0)}(\bar{r}, \bar{r}_1)$ is the dyadic Green's function, and $\overline{E}_1^{(0)}$ the zeroth-order solution in the absence of ϵ_{1f}, namely, $\epsilon_{1f} = 0$.

The solution for the electric field in region 0 is

$$\overline{E}_0 = \overline{E}_0^{(0)} + \iiint_{V_1} d^3\bar{r}_1 \, \overline{\overline{G}}_{01}^{(0)}(\bar{r}, \bar{r}_1) \cdot Q(\bar{r}_1) \overline{E}_1(\bar{r}_1) \tag{4}$$

The first term in (4) is the zeroth-order solution $\overline{E}_0$ in the absence of ϵ_{1f}, representing the specularly reflected wave, which is also called the coherent component of the total field. The second term is the scattered field $\overline{E}_s$

$$\overline{E}_s(\bar{r}) = \iiint_{V_1} d^3\bar{r}_1 \, \overline{\overline{G}}_{01}(\bar{r}, \bar{r}_1) \cdot Q(\bar{r}_1) \overline{E}_1(\bar{r}_1) \tag{5}$$

Equation (5) can be solved with an iterative approach. We assume that the total field can be expanded in terms of a Born series

$$\overline{E}_l = \sum_{n=0}^{\infty} \overline{E}_l^{(n)} \qquad l = 0, 1 \tag{6}$$

Substituting in (4), we find that the nth-order $\overline{E}$ field is determined by the $(n-1)$th-order field,

$$\overline{E}_s^{(n)} = \iiint_{V_1} d^3\bar{r}_1 \, \overline{\overline{G}}_{01}(\bar{r},\bar{r}_1) \cdot Q(\bar{r}_1)\overline{E}_1^{(n-1)}(\bar{r}_1) \tag{7}$$

Notice that $\overline{E}_i^{(n)}$ with $n \neq 0$ are all randomly fluctuating fields.

Forming the absolute square of $\overline{E}_s^{(1)}$ and taking ensemble average, we obtain the first-order scattered intensity

$$< |\overline{E}_s^{(1)}(\bar{r})|^2 > = \iiint_{V_1} d^3\bar{r}_1 \iiint_{V_1} d^3\bar{r}_2 \, \overline{\overline{G}}_{01}^{(0)}(\bar{r},\bar{r}_1) \cdot \overline{E}_1^{(0)}(\bar{r}_1)$$

$$\cdot \overline{\overline{G}}_{01}^{(0)*}(\bar{r},\bar{r}_2) \cdot \overline{E}_1^{(0)*}(\bar{r}_2) < Q(\bar{r}_1)Q^*(\bar{r}_2) > \tag{8}$$

where $< Q(\bar{r}_1)Q^*(\bar{r}_2) >$ is a two-point correlation function. For a statistically homogeneous medium, the correlation function depends only upon the separation of the points $\bar{r}_1$ and $\bar{r}_2$, that is,

$$< Q(\bar{r}_1)Q^*(\bar{r}_2) > = C(\bar{r}_1 - \bar{r}_2) \tag{9}$$

Consider the correlation function

$$C(\bar{r}_1 - \bar{r}_2) = \delta |k_1|^4 e^{-|\bar{r}_1 - \bar{r}_2|/r_0} \tag{10}$$

where δ is the variance and r_0 the correlation length of the permittivity fluctuations. The variance is related to the strength of fluctuations and the correlation length corresponds roughly to the sizes of the scatterers. We see that major contributions to the scattered-field intensity come from fluctuations separated by no more than a correlation length.

a. Dyadic Green's Function for Layered Media

Consider a layered medium with a source located in region 0. In the absence of the layered medium, the dyadic Green's function is governed by

$$\nabla \times \nabla \times \overline{\overline{G}}(\bar{r},\bar{r}') - k^2\overline{\overline{G}}(\bar{r},\bar{r}') = \overline{\overline{I}}\delta(\bar{r} - \bar{r}') \tag{11}$$

where we have determined in unbounded space

$$\overline{\overline{G}}(\bar{r},\bar{r}') = \left[\overline{\overline{I}} + \frac{1}{k^2}\nabla\nabla\right] g(\bar{r},\bar{r}') \tag{12}$$

$$g(\bar{r},\bar{r}') = \frac{e^{ik(\bar{r}-\bar{r}')}}{4\pi|\bar{r} - \bar{r}'|} \tag{13}$$

Let the source be placed at the origin, $\bar{r}' = 0$. We have

$$(\nabla^2 + k^2)g(\bar{r}) = -\delta(\bar{r}) \tag{14}$$

Fourier transformation gives

$$\delta(\bar{r}) = \frac{1}{(2\pi)^3}\iiint dk_x\, dk_y\, dk_z\, e^{i\bar{k}\cdot\bar{r}} \tag{15}$$

$$g(\bar{r}) = \frac{1}{(2\pi)^3}\iiint dk_x\, dk_y\, dk_z\, e^{i\bar{k}\cdot\bar{r}}\, g(\bar{k}) \tag{16}$$

with

$$g(\bar{k}) = \frac{1}{k_x^2 + k_y^2 + k_z^2 - k^2} \tag{17}$$

as derived from (14).

For the triple integral in (16), we integrate over k_z by noting from (17) that poles occur at $k_z^2 = k^2 - k_x^2 - k_y^2$. For $z > 0$, we deform the contour upward such that $Im\{k_z\} > 0$. For $z < 0$, we deform the contour downward. We obtain

$$g(\bar{r}) = \begin{cases} \dfrac{i}{(2\pi)^2}\iint dk_x\, dk_y\, \dfrac{1}{2k_{0z}} e^{ik_x x + ik_y y + ik_{0z} z} & z > 0 \\ \dfrac{i}{(2\pi)^2}\iint dk_x\, dk_y\, \dfrac{1}{2k_{0z}} e^{ik_x x + ik_y y - ik_{0z} z} & z < 0 \end{cases} \tag{18}$$

where

$$k_{0z} = \sqrt{k^2 - k_x^2 - k_y^2} \tag{19}$$

To find the expression for $\overline{\overline{G}}(\bar{r})$, we make use of (12) and notice that there is a discontinuity in $\partial g(\bar{r})/\partial z$ at $z = 0$, which gives

$$\frac{\partial^2}{\partial z^2} g(\bar{r}) = -\delta(\bar{r}) - \begin{cases} \dfrac{i}{(2\pi)^2}\iint dk_x\, dk_y\, \dfrac{k_{0z}}{2} e^{ik_x x + ik_y y + ik_{0z} z} & z > 0 \\ \dfrac{i}{(2\pi)^2}\iint dk_x\, dk_y\, \dfrac{k_{0z}}{2} e^{ik_x x + ik_y y - ik_{0z} z} & z < 0 \end{cases} \tag{20}$$

We find the expression for $\overline{\overline{G}}(\bar{r})$ to be

$$\overline{\overline{G}}(\bar{r}) = -\hat{z}\hat{z}\frac{1}{k^2}\delta(\bar{r}) + \begin{cases} \dfrac{i}{8\pi^2}\iint dk_x\, dk_y\, \dfrac{1}{k_{0z}}\left[\overline{\overline{I}} - \dfrac{\bar{k}\,\bar{k}}{k^2}\right] e^{i\bar{k}\cdot\bar{r}} & z > 0 \\ \dfrac{i}{8\pi^2}\iint dk_x\, dk_y\, \dfrac{1}{k_{0z}}\left[\overline{\overline{I}} - \dfrac{\bar{K}\,\bar{K}}{k^2}\right] e^{i\bar{K}\cdot\bar{r}} & z < 0 \end{cases} \tag{21}$$

where

$$\bar{k} = \hat{x}k_x + \hat{y}k_y + \hat{z}k_{0z} \tag{22}$$

$$\bar{K} = \hat{x}k_x + \hat{y}k_y - \hat{z}k_{0z} \tag{23}$$

Recognizing that $\hat{k} = \bar{k}/k$, we form an orthonormal system consisting of unit vectors $\hat{k}$, $\hat{h}(k_{0z})$, and $\hat{e}(k_{0z})$ as follows:

$$\hat{e}(k_{0z}) = \frac{\hat{k} \times \hat{z}}{|\hat{k} \times \hat{z}|} = \frac{(\hat{x}k_y - \hat{y}k_x)}{\sqrt{k_x^2 + k_y^2}} \tag{24}$$

$$\hat{h}(k_{0z}) = \frac{1}{k}\hat{e} \times \bar{k} = \frac{-k_{0z}}{k\sqrt{k_x^2 + k_y^2}}(\hat{x}k_x + \hat{y}k_y) + \hat{z}\frac{\sqrt{k_x^2 + k_y^2}}{k} \tag{25}$$

As $\bar{\bar{I}} = \hat{k}\hat{k} + \hat{e}\hat{e} + \hat{h}\hat{h}$, we have $\bar{\bar{I}} - \hat{k}\hat{k} = \hat{e}\hat{e} + \hat{h}\hat{h}$. The dyadic Green's function $\bar{\bar{G}}(\bar{r}, \bar{r}')$, after translating the origin to $\bar{r}'$, becomes

$$\begin{aligned}\bar{\bar{G}}(\bar{r}, \bar{r}') &= -\hat{z}\hat{z}\frac{1}{k^2}\delta(\bar{r} - \bar{r}') \\ &+ \begin{cases} \dfrac{i}{8\pi^2}\displaystyle\iint dk_x\, dk_y\, \frac{1}{k_{0z}} \Big\{ \Big[\hat{e}(k_{0z})e^{i\bar{k}\cdot\bar{r}}\Big]\, \hat{e}(k_{0z})e^{-i\bar{k}\cdot\bar{r}'} \\ \qquad + \Big[\hat{h}(k_{0z})e^{i\bar{k}\cdot r}\Big]\, \hat{h}(k_{0z})e^{-i\bar{k}\cdot\bar{r}'} \Big\} & z > z' \\ \dfrac{i}{8\pi^2}\displaystyle\iint dk_x\, dk_y\, \frac{1}{k_{0z}} \Big\{ \Big[\hat{e}(-k_{0z})e^{i\bar{K}\cdot\bar{r}}\Big]\, \hat{e}(-k_{0z})e^{-i\bar{K}\cdot\bar{r}'} \\ \qquad + \Big[\hat{h}(-k_{0z})e^{i\bar{K}\cdot r}\Big]\, \hat{h}(-k_{0z})e^{-i\bar{K}\cdot\bar{r}'} \Big\} & z < z' \end{cases}\end{aligned} \tag{26}$$

where $\hat{K} = \bar{K}/k$, $\hat{e}(-k_{0z}) = \hat{e}(k_{0z})$, and $\hat{h}(-k_{0z}) = \hat{e} \times \bar{K}/k$ form another orthonormal set of unit vectors about the wave vector $\bar{K}$.

To derive the dyadic Green's function $\bar{\bar{G}}_{l0}(\bar{r}, \bar{r}')$ for a layered medium, we use the first subscript l to denote the region of the observation point and the second subscript 0 to indicate that the source is in region 0. To facilitate the matching of the boundary conditions, we consider, for the Green's function $\bar{\bar{G}}(\bar{r}, \bar{r}')$, only $z < z'$ and thus neglect the delta function term. We write

$$\overline{\overline{G}}_{00}(\bar{r},\bar{r}') = \frac{i}{8\pi^2}\iint dk_x\,dk_y\,\frac{1}{k_{0z}}\Big\{\Big[\hat{e}(-k_{0z})e^{i\overline{K}\cdot\bar{r}} + R^{\mathrm{TE}}\hat{e}(k_{0z})e^{i\bar{k}\cdot\bar{r}}\Big]\hat{e}(-k_{0z})e^{-i\overline{K}\cdot\bar{r}'} + \Big[R^{\mathrm{TM}}\hat{h}(k_{0z})e^{i\bar{k}\cdot\bar{r}} + \hat{h}(-k_{0z})e^{i\overline{K}\cdot\bar{r}}\Big]\hat{h}(-k_{0z})e^{-i\overline{K}\cdot\bar{r}'}\Big\} \qquad \text{for } z < z' \tag{27}$$

$$\overline{\overline{G}}_{l0}(\bar{r},\bar{r}') = \frac{i}{8\pi^2}\iint dk_x\,dk_y\,\frac{1}{k_{0z}}\Big\{\Big[B_l\hat{e}_l(-k_{lz})e^{i\overline{K}_l\cdot\bar{r}} + A_l\hat{e}_l(k_{lz})e^{i\bar{k}_l\cdot\bar{r}}\Big]\hat{e}(-k_{0z})e^{-i\overline{K}\cdot\bar{r}'} + \Big[C_l\hat{h}_l(k_{lz})e^{i\bar{k}_l\cdot\bar{r}} + D_l\hat{h}_l(-k_{lz})e^{i\overline{K}_l\cdot\bar{r}}\Big]\hat{h}(-k_{0z})e^{-i\overline{K}\cdot\bar{r}'}\Big\} \tag{28}$$

$$\overline{\overline{G}}_{t0}(\bar{r},\bar{r}') = \frac{i}{8\pi^2}\iint dk_x\,dk_y\,\frac{1}{k_{0z}}\Big\{T^{\mathrm{TE}}\,\hat{e}_t(-k_{tz})e^{i\overline{K}_t\cdot\bar{r}}\,\hat{e}(-k_{0z})e^{-i\overline{K}\cdot\bar{r}'} + T^{\mathrm{TM}}\,\hat{h}_t(-k_{tz})e^{i\overline{K}_t\cdot\bar{r}}\,\hat{h}(-k_{0z})e^{-i\overline{K}\cdot\bar{r}'}\Big\} \tag{29}$$

where

$$k_{lz} = \sqrt{k_l^2 - k_x^2 - k_y^2} \tag{30}$$

$$\bar{k}_l = \hat{x}k_x + \hat{y}k_y + \hat{z}k_{lz} \tag{31}$$

$$\overline{K}_l = \hat{x}k_x + \hat{y}k_y - \hat{z}k_{lz} \tag{32}$$

The coefficients A_l, B_l, C_l and D_l are then determined by the boundary conditions.

To simplify the algebra, we consider a two-layer medium where $t = 2$. The procedure will be easily generalized to that for a general layered medium. With boundaries at $z = 0$ and $z = -d$, we impose the boundary conditions of continuity of $\hat{z} \times \overline{\overline{G}}$ and $\hat{z} \times \nabla \times \overline{\overline{G}}/\mu$ corresponding to continuity of the tangential $\overline{E}$ and $\overline{H}$ fields. We find

$$R^{\mathrm{TE}} + 1 = A_1 + B_1 \tag{33}$$

$$\frac{k_{0z}}{k}(R^{\mathrm{TM}} - 1) = \frac{k_{1z}}{k_1}(C_1 - D_1) \tag{34}$$

$$k_{0z}(R^{\mathrm{TE}} - 1) = k_{1z}(A_1 - B_1) \tag{35}$$

$$k(R^{\mathrm{TM}} + 1) = k_1(C_1 + D_1) \tag{36}$$

and

$$A_1 e^{-ik_{1z}d} + B_1 e^{ik_{1z}d} = T^{\mathrm{TE}} e^{ik_{2z}d} \tag{37}$$

$$\frac{k_{1z}}{k_1}(C_1 e^{-ik_{1z}d} - D_1 e^{ik_{1z}d}) = -\frac{k_{2z}}{k_2} T^{\mathrm{TM}} e^{ik_{2z}d} \tag{38}$$

$$k_{1z}(A_1 e^{-ik_{1z}d} - B_1 e^{ik_{1z}d}) = -k_{2z} T^{\mathrm{TE}} e^{ik_{2z}d} \tag{39}$$

$$k_1(C_1 e^{-ik_{1z}d} + D_1 e^{ik_{1z}d}) = k_2 T^{\mathrm{TM}} e^{ik_{2z}d} \tag{40}$$

Solving (33)–(40), we obtain all eight unknown wave amplitudes. In particular,

$$\begin{aligned}\overline{\overline{G}}_{10}(\bar{r},\bar{r}') = \frac{i}{8\pi}\iint dk_x\, dk_y\, \frac{1}{k_{1z}} \\ \left\{ \frac{T_{10}^{\mathrm{TE}}}{D_2(k_\perp)} \left[R_{12}^{\mathrm{TE}} e^{i2k_{1z}d} \hat{e}_1(k_{1z}) e^{i\bar{k}_1\cdot\bar{r}} + \hat{e}_1(-k_{1z}) e^{i\bar{K}_1\cdot\bar{r}} \right] \hat{e}(-k_{0z}) e^{-i\bar{K}\cdot\bar{r}'} \right. \\ \left. + \frac{k_1}{k} \frac{T_{10}^{\mathrm{TM}}}{F_2(k_\perp)} \left[R_{12}^{\mathrm{TM}} e^{i2k_{1z}d} \hat{h}_1(k_{1z}) e^{i\bar{k}_1\cdot\bar{r}} + \hat{h}_1(-k_{1z}) e^{i\bar{K}_1\cdot\bar{r}} \right] \hat{h}(-k_{0z}) e^{-i\bar{K}\cdot\bar{r}'} \right\}\end{aligned} \tag{41}$$

where

$$T_{10}^{\mathrm{TE}} = 1 + R_{10}^{\mathrm{TE}} = \frac{2k_{1z}}{k_z + k_{1z}} \tag{42}$$

$$T_{10}^{\mathrm{TM}} = 1 + R_{10}^{\mathrm{TM}} = \frac{2\epsilon k_{1z}}{\epsilon_1 k_z + \epsilon k_{1z}} \tag{43}$$

$$D_2(k_\perp) = 1 + R_{01}^{\mathrm{TE}} R_{12}^{\mathrm{TE}} e^{i2k_{1z}d} \tag{44}$$

$$F_2(k_\perp) = 1 + R_{01}^{\mathrm{TM}} R_{12}^{\mathrm{TM}} e^{i2k_{1z}d} \tag{45}$$

$$R_{12}^{\mathrm{TE}} = \frac{k_{1z} - k_{2z}}{k_{1z} + k_{2z}} \tag{46}$$

$$R_{12}^{\mathrm{TM}} = \frac{\epsilon_2 k_{1z} - \epsilon_1 k_{2z}}{\epsilon_2 k_{1z} + \epsilon_1 k_{2z}} \tag{47}$$

For latter applications we shall be interested in $\overline{\overline{G}}_{01}(\bar{r},\bar{r}')$ for observation point in region 0 and source in region 1. The symmetric property for dyadic Green's functions calls for [Tai, 1971]

$$\overline{\overline{G}}_{01}(\bar{r},\bar{r}') = \overline{\overline{G}}_{10}^{T}(\bar{r}',\bar{r}) \tag{48}$$

We transpose $\overline{\overline{G}}_{10}$ and change $k_x \to -k_x$, and $k_y = -k_y$. From (24)–(25) we also have $\hat{e}(k_{0z}) \to \hat{e}(-k_{0z})$ and $\hat{h}(k_{0z}) \to \hat{h}(-k_{0z})$. Equation (41) then gives

$$\overline{\overline{G}}_{01}(\bar{r}, \bar{r}') = \iint d\bar{k}_\perp \, \overline{\overline{g}}_{01}(\bar{k}_\perp, z, z') e^{i\bar{k}_\perp \cdot (\bar{r}_\perp - \bar{r}'_\perp)} \tag{49}$$

where

$$\begin{aligned}
\overline{\overline{g}}_{01}(\bar{k}_\perp, z, z') &= \frac{i}{8\pi^2} \frac{e^{ik_z z}}{k_{1z}} \\
&\left\{ \frac{T_{10}^{\mathrm{TE}}}{D_2(k_\perp)} \hat{e}(k_{0z}) \left[R_{12}^{\mathrm{TE}} e^{i2k_{1z}d} \hat{e}_1(-k_{1z}) e^{ik_{1z}z'} + \hat{e}_1(k_{1z}) e^{-ik_{1z}z'} \right] \right. \\
&\left. + \frac{k_1}{k} \frac{T_{10}^{\mathrm{TM}}}{F_2(k_\perp)} \hat{h}(k_{0z}) \left[R_{12}^{\mathrm{TM}} e^{i2k_{1z}d} \hat{h}_1(-k_{1z}) e^{ik_{1z}z'} + \hat{h}_1(k_{1z}) e^{-ik_{1z}z'} \right] \right\}
\end{aligned} \tag{50}$$

The integration in (49) can be evaluated for the radiation field with a two-dimensional stationary-phase method [Born and Wolf, 1975]. As $kr \to \infty$, the dominant term in (49) is

$$e^{i\bar{k}\cdot\bar{r}} = \exp\left[i(k_x x + k_y y + \sqrt{k^2 - k_x^2 - k_y^2}\, z) \right]$$

With the observation point at $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, and $z = r\cos\theta$, the stationary-phase points are found to be

$$\begin{aligned}
k_x &= k\sin\theta\cos\phi \\
k_y &= k\sin\theta\sin\phi
\end{aligned}$$

The dyadic Green's function $\overline{\overline{G}}_{01}(\bar{r}, \bar{r}')$ is then determined. We obtain

$$\begin{aligned}
\overline{\overline{G}}_{01}(\bar{r}, \bar{r}') = \frac{e^{ikr}}{4\pi r} \left\{ \left[\frac{T_{01}^{\mathrm{TE}}}{D_2} \hat{e}(k_{0z}) \hat{e}_1(k_{1z}) + \frac{k}{k_1} \frac{T_{01}^{\mathrm{TM}}}{F_2} \hat{h}(k_{0z}) \hat{h}_1(k_{1z}) \right] e^{-i\bar{k}_1 \cdot \bar{r}'} \right. \\
+ \left[\frac{T_{01}^{\mathrm{TE}}}{D_2} R_{12}^{\mathrm{TE}} \hat{e}(k_{0z}) \hat{e}_1(-k_{1z}) + \frac{k}{k_1} \frac{T_{01}^{\mathrm{TM}}}{F_2} R_{12}^{\mathrm{TM}} \hat{h}(k_{0z}) \hat{h}_1(-k_{1z}) \right] \\
\left. \cdot e^{i2k_{1z}d} e^{-i\bar{K}_1 \cdot \bar{r}'} \right\}
\end{aligned} \tag{51}$$

with all k_x and k_y equal to the values at the stationary-phase point.

b. Scattering by a Half-Space Random Medium

To illustrate with a half-space random medium, we consider scattered fields in the far-field zone. Let an incident wave be horizontally polarized with

$$\bar{E}_{0i} = \hat{e}(-k_{0zi})E_0\, e^{-ik_{0zi}z} e^{i\bar{k}_{\perp i}\cdot\bar{r}}$$

where $k_{0zi} = k\cos\theta_i$. The unperturbed field in region 1 is

$$\bar{E}_1^{(0)}(\bar{r}) = E_0\, T_{01i}^{\mathrm{TE}}\, \hat{e}_{1i}(k_{1zi}) e^{i\bar{k}_{1i}\cdot\bar{r}}$$

where $k_{1zi} = (k_1^2 - k^2\sin^2\theta_i)^{1/2}$ and $\bar{k}_{1i} = \bar{k}_{\perp i} - \hat{z}k_{1zi}$. The subscript i in T_{01i}^{TE} signifies the fact that the k_z values are those in the incident direction. Making use of the far-field dyadic Green's function derived in the last section and simplify to that for a half-space medium, we have

$$\bar{\bar{G}}_{01}(\bar{r},\bar{r}') = \frac{e^{ikr}}{4\pi r}\left\{T_{01s}^{\mathrm{TE}}\hat{e}(k_{0zs})\hat{e}_1(k_{1zs}) + \frac{k}{k_1}T_{01s}^{\mathrm{TM}}\hat{h}(k_{0zs})\hat{h}_1(k_{1zs})\right\} e^{-i\bar{k}_{1s}\cdot\bar{r}'}$$

with

$$k_{0zs} = \sqrt{k^2 - k_{xs}^2 - k_{ys}^2}$$

$$k_{1zs} = \sqrt{k_1^2 - k_{xs}^2 - k_{ys}^2}$$

$$\bar{k}_{1s} = \bar{k}_{\perp s} + \hat{z}k_{1zs}$$

The first-order scattered field intensity is obtained as follows:

$$\langle|\bar{E}_s^{(1)}(\bar{r})|^2\rangle = \frac{|E_0|^2}{16\pi^2 r^2}\left|\left[T_{01s}^{\mathrm{TE}}\,\hat{e}(k_{0zs})\hat{e}_1(k_{1zs}) + \frac{k}{k_1}T_{01s}^{\mathrm{TM}}\,\hat{h}(k_{0zs})\hat{h}_1(k_{1zs})\right]\cdot T_{01i}^{\mathrm{TE}}\,\hat{e}_{1i}(k_{1zi})\right|^2 W \quad (52)$$

where

$$W = \iiint_{V_1} d^3\bar{r}_1 \iiint_{V_1} d^3\bar{r}_2\; C(\bar{r}_1 - \bar{r}_2) e^{i(\bar{k}_{1i}-\bar{k}_{1s})\cdot\bar{r}_1 - i(\bar{k}_{1i}^*-\bar{k}_{1s}^*)\cdot\bar{r}_2} \quad (53)$$

The correlation function may be expressed as the Fourier transform of a spectral intensity function Φ

$$C(\bar{r}_1 - \bar{r}_2) = \delta|k_1|^4 \iiint_{-\infty}^{\infty} d\bar{k}\; \Phi(\bar{k}) e^{-i\bar{k}\cdot(\bar{r}_1-\bar{r}_2)} \quad (54)$$

For the correlation function in (10), for instance, we choose the coordinates with the z axis along $\bar{k}$ without loss of generality and find

$$\begin{aligned}\Phi(\bar{k}) &= \frac{1}{8\pi^3}\iiint d^3\bar{r}\; e^{-r/r_0} e^{i\bar{k}\cdot\bar{r}} \\ &= \frac{1}{4\pi^2}\int_0^\infty dr\; r^2 e^{-r/r_0}\int_0^\pi d\theta\; \sin\theta\, e^{ikr\cos\theta} \\ &= \frac{r_0^3}{\pi^2(1+k^2r_0^2)^2}\end{aligned}$$

Substitution of (54) into (53) yields

$$\begin{aligned}W &= 4\pi^2\delta|k_1|^4 \iiint_{-\infty}^{\infty} d^3\bar{k} \iiint dx_1 dy_1 dz_1\, \exp(i(\bar{k}_{1zi} - \bar{k}_{1zs} - \bar{k}_z)\cdot z_1) \\ &\cdot \exp(i(k_{1zi}^* + k_{1zs}^* + k_z)z_2)\cdot\Phi(\bar{k})\delta(k_{1xi} - k_{1xs} - k_x, k_{1yi} - k_{1ys} - k_y) \\ &= 4\pi^2\delta|k_1|^4\int_{-\infty}^{\infty} dk_z \int_{-\infty}^{0} dz_1 \int_{-\infty}^{0} dz_2\; \Phi(k_{xi} - k_{xs}, k_{yi} - k_{ys}, k_z) \\ &\quad\cdot \exp(-i(k_{1zi} + k_{1zs} + k_z)z_1)\exp(i(k_{1zi}^* + k_{1zs}^* + k_z)z_2) \\ &= \delta|k_1|^4 A\int_{-\infty}^{\infty} dk_z \frac{\Phi(k_{xi} - k_{xs}, k_{yi} - k_{ys}, k_z)}{(k_{1zi} + k_{1zs} + k_z)(k_{1zi}^* + k_{1zs}^* + k_z)}\end{aligned}$$

where A is the illuminated area resulting from the integration over $dx_1\, dy_1$. Consider low absorption with $k_{1I} \ll k_{1R}$. We apply contour integration and recognize that most of the contribution of the integral comes from the residue of the pole at $k_z = -k_{1zi}^* - k_{1zs}^*$. We find

$$W \approx 4\pi^3\delta|k_1|^4 A\frac{\Phi(k_{xi} - k_{xs}, k_{yi} - k_{ys}, -k_{1zi}^* - k_{1zs}^*)}{k_{1ziI} + k_{1zsI}}$$

The bistatic scattering coefficients $\gamma_{\mu\nu}(\hat{k}_s, \hat{k}_i)$ are defined as

$$\gamma_{\mu\nu} = \frac{4\pi r^2 < |\bar{E}_s|^2 >_\nu}{A\cos\theta_i |\bar{E}_0|_\mu^2}$$

where μ denotes the incident polarization and ν the scattered polar-

ization. We obtain, for the case $\phi_i = 0$,

$$\begin{Bmatrix} \gamma_{hh} \\ \gamma_{vh} \\ \gamma_{hv} \\ \gamma_{vv} \end{Bmatrix} = \frac{\delta |k_1|^4 \pi^2 \Phi}{\cos\theta_i (k_{1zI} + k_{1zsI})} \cdot \begin{Bmatrix} |T_{01s}^{\mathrm{TE}} T_{01i}^{\mathrm{TE}}|^2 \cos^2\phi_s \\ |k_{1zs} \frac{k}{k_1} T_{01s}^{\mathrm{TM}} T_{01i}^{\mathrm{TE}}|^2 \sin^2\phi_s \\ |k_{1zi} \frac{k}{k_1} T_{01s}^{\mathrm{TE}} T_{01i}^{\mathrm{TM}}|^2 \sin^2\phi_s \\ |T_{01s}^{\mathrm{TM}} T_{01i}^{\mathrm{TM}}|^2 \frac{k^6 \sin^2\theta_s}{k_1^8} |k_{xi} - \frac{k_{xs} k_{1zs} k_{1zi}}{k_{xs}^2 + k_{ys}^2}|^2 \end{Bmatrix}$$

For the backscattering coefficient

$$\sigma_{\mu\nu} = \gamma_{\mu\nu}(\theta_s = \theta_i,\ \phi_s = \pi + \phi_i) \cos\theta_i$$

We see that $\sigma_{\mu\nu} = \sigma_{h\nu} = 0$ for the first-order solution from the Born series. To determine the cross-polarized backscattering intensities, we must carry out the Born approximation to the second-order from the iterative procedure.

The above results are obtained with the Born approximation that makes use of the unperturbed field values for the background medium in the absence of the inhomogeneities. In the distorted Born approximation, which often yields better numerical results as compared with experimental measurements, the unperturbed field values in the background medium with the effective permittivity are used. The Born approximations are valid only for weak permittivity fluctuations. When the variance of the randomly fluctuating part is large, we must resort to the strong fluctuation theory in which the singularities of the dyadic Green's functions are properly taken care of and a new small parameter such as ξ in 6.6 can be utilized to effect approximations similar to the Born series.

6.6 Scattering by Random Rough Surfaces

Two analytical approaches have been applied to the study of the scattering of electromagnetic waves by random rough surfaces. In the

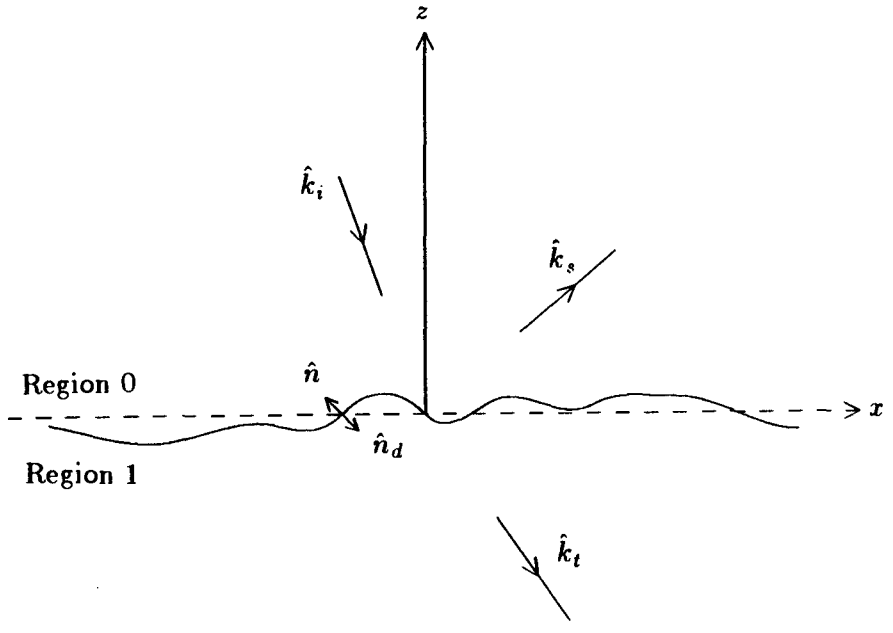


Figure 6.6.1 Scattering by a random rough surface.

Kirchhoff approximation (KA), the fields at any point on the surface are approximated by the fields that would be present on the tangent plane at that point. Thus the tangent plane approximation requires a large radius of curvature relative to the incident wavelength at every point on the surface. In the small perturbation method (SPM) the surface variations are assumed to be much smaller than the incident wavelength and the slopes of the rough surface are relatively small.

Consider a plane wave incident upon a random rough surface [Fig. 6.6.1]. The electric field of the incident wave is given by

$$\overline{E}_i = \hat{e}_i \, E_o \, e^{i\overline{k}_i \cdot \overline{r}}$$

where $\overline{k}_i$ is the incident wave vector and $\hat{e}_i$ the polarization of the electric field vector. The rough surface is characterized by a random height distribution $z = f(\overline{r}_\perp)$ where $f(\overline{r}_\perp)$ is a Gaussian random variable with zero mean, $< f(\overline{r}_\perp) >= 0$. From Huygens' principle, which expresses the field at an observation point in terms of fields at the boundary surface, the following expressions are obtained for the

scattered fields in region 0 and the transmitted fields in region 1:

$$\overline{E}_s(\bar{r}) = \iint_{S'} dS' \left\{ i\omega\mu_o \overline{\overline{G}}(\bar{r},\bar{r}') \cdot [\hat{n} \times \overline{H}(\bar{r}')] + \nabla \times \overline{\overline{G}}(\bar{r},\bar{r}') \cdot [\hat{n} \times \overline{E}(\bar{r}')] \right\} \tag{1a}$$

$$\overline{E}_t(\bar{r}) = \iint_{S'} dS' \left\{ i\omega\mu_o \overline{\overline{G}}_1(\bar{r},\bar{r}') \cdot [\hat{n}_d \times \overline{H}(\bar{r}')] + \nabla \times \overline{\overline{G}}_1(\bar{r},\bar{r}') \cdot [\hat{n}_d \times \overline{E}(\bar{r}')] \right\} \tag{1b}$$

where S' denotes the rough surface on which the surface integration is to be carried out, $\hat{n}$ and $\hat{n}_d$ are the unit vectors normal to the rough surface and pointing into the reflected and transmitted regions [Fig. 6.6.1]. The dyadic Green's function for the homogeneous space of regions 0 and 1, $\overline{\overline{G}}(\bar{r},\bar{r}')$ and $\overline{\overline{G}}_1(\bar{r},\bar{r}')$, are

$$\overline{\overline{G}}(\bar{r},\bar{r}') = \left[\overline{\overline{I}} + \frac{1}{k^2}\nabla\nabla\right] \frac{e^{ik|\bar{r}-\bar{r}'|}}{4\pi|\bar{r}-\bar{r}'|} \tag{2a}$$

and

$$\overline{\overline{G}}_1(\bar{r},\bar{r}') = \left[\overline{\overline{I}} + \frac{1}{k_1^2}\nabla\nabla\right] \frac{e^{ik_1|\bar{r}-\bar{r}'|}}{4\pi|\bar{r}-\bar{r}'|} \tag{2b}$$

where $k = \omega\sqrt{\mu_o\epsilon_o}$ and $k_1 = \omega\sqrt{\mu_o\epsilon_1}$. If the observation point is in the far field region, then the dyadic Green's functions simplify to

$$\overline{\overline{G}}(\bar{r},\bar{r}') \simeq (\overline{\overline{I}} - \hat{k}_s\hat{k}_s) \frac{e^{ikr}}{4\pi r} \exp(-i\bar{k}_s \cdot \bar{r}') \tag{3}$$

$$\overline{\overline{G}}_1(\bar{r},\bar{r}') \simeq (\overline{\overline{I}} - \hat{k}_t\hat{k}_t) \frac{e^{ik_1 r}}{4\pi r} \exp(-i\bar{k}_t \cdot \bar{r}') \tag{4}$$

where $\hat{k}_s$ and $\hat{k}_t$ denote the scattered and the transmitted directions in regions 0 and 1.

Substituting (3) and (4) into the diffraction integral (1), we obtain, in the reflected direction $\hat{k}_s$ and the transmitted direction $\hat{k}_t$,

$$\overline{E}_s(\bar{r}) = \frac{ik\,e^{ikr}}{4\pi r}(\overline{\overline{I}} - \hat{k}_s\hat{k}_s) \cdot \iint_{S'} dS' \left\{ \hat{k}_s \times [\hat{n} \times \overline{E}(\bar{r}')] + \eta[\hat{n} \times \overline{H}(\bar{r}')] \right\} e^{-i\bar{k}_s \cdot \bar{r}'} \tag{5a}$$

$$\overline{E}_t(\bar{r}) = \frac{ik_1 e^{ik_1 r}}{4\pi r}(\overline{\overline{I}} - \hat{k}_t\hat{k}_t) \cdot \iint_{S'} dS' \left\{\hat{k}_t \times [\hat{n}_d \times \overline{E}(\bar{r}')] + \eta_1[\hat{n}_d \times \overline{H}(\bar{r}')]\right\} e^{-i\bar{k}_t\cdot\bar{r}'} \quad (5b)$$

where η and η_1 are the wave impedances in regions 0 and 1.

a. Kirchhoff Approximation

We first form an orthonormal system $(\hat{p}_i, \hat{q}_i, \hat{k}_i)$ at a point $\bar{r}'$, with

$$\hat{q}_i = \frac{\hat{k}_i \times \hat{n}}{|\hat{k}_i \times \hat{n}|} \quad (6a)$$

$$\hat{p}_i = \hat{q}_i \times \hat{k}_i \quad (6b)$$

where, $\hat{n}(\bar{r}') = -\hat{n}_d(\bar{r}')$, is the normal to the surface at the point $\bar{r}'$ pointing into region 0. The unit vectors are the local perpendicular and parallel polarization vectors at the point $\bar{r}'$. In applying the tangent plane approximation, we solve the boundary value problem for the TE and TM polarization of a wave incident onto an infinite planar interface with the tangent plane as the interface. The incident field is decomposed into locally perpendicular and parallel polarization fields.

The TE component of the incident field is $(\hat{e}_i \cdot \hat{q}_i)\hat{q}_i E_o e^{i\bar{k}_i\cdot\bar{r}'}$ and the local reflected field is $R^{\mathrm{TE}}(\hat{e}_i \cdot \hat{q}_i)\hat{q}_i E_o e^{i\bar{k}_i\cdot\bar{r}'}$ where R^{TE} is the local Fresnel reflection coefficient for the TE component

$$R^{\mathrm{TE}} = \frac{k\cos\theta_{l_i} - \sqrt{k_1^2 - k^2\sin^2\theta_{l_i}}}{k\cos\theta_{l_i} + \sqrt{k_1^2 - k^2\sin^2\theta_{l_i}}}$$

where θ_{l_i} is the local angle of incidence at the point $\bar{r}'$. The magnetic fields associated with the incident and reflected fields are $\hat{k}_i \times (\hat{e}_i \cdot \hat{q}_i)\hat{q}_i E_o e^{i\bar{k}_i\cdot\bar{r}'}/\eta$ and $R^{\mathrm{TE}}\hat{k}_r \times (\hat{e}_i \cdot \hat{q}_i)\hat{q}_i E_o e^{i\bar{k}_i\cdot\bar{r}'}/\eta$ where $\hat{k}_r$ is the local reflected direction and is related to the incident direction by

$$\hat{k}_r = \hat{k}_i - 2\hat{n}(\hat{n}\cdot\hat{k}_i) \quad (7)$$

Hence, the tangential electric field of this perpendicular component at the point $\bar{r}'$ is

$$\hat{n} \times \overline{E} = (\hat{n} \times \hat{q}_i)(\hat{e}_i \cdot \hat{q}_i)(1 + R^{\mathrm{TE}}) E_o e^{i\bar{k}_i\cdot\bar{r}'} \quad (8a)$$

and the associated magnetic field is

$$\hat{n} \times \overline{H} = \frac{1}{\eta}(\hat{e}_i \cdot \hat{q}_i)\hat{n} \times \left[(\hat{k}_i \times \hat{q}_i) + R^{\mathrm{TE}}(\hat{k}_r \times \hat{q}_i)\right] E_o\, e^{i\overline{k}_i \cdot \overline{r}'}$$

$$= -(1 - R^{\mathrm{TE}})(\hat{n} \cdot \hat{k}_i)\frac{(\hat{e}_i \cdot \hat{q}_i)}{\eta}\hat{q}_i\, E_o\, e^{i\overline{k}_i \cdot \overline{r}'} \tag{8b}$$

where we have made use of the relations $\hat{n} \cdot \hat{q}_i = 0$ and $\hat{n} \cdot \hat{k}_r = -\hat{n} \cdot \hat{k}_i$. The calculations can be repeated for the local TM component with local reflection coefficient

$$R^{\mathrm{TM}} = \frac{\epsilon_1 k \cos\theta_{l_i} - \epsilon_o\sqrt{k_1^2 - k^2 \sin^2\theta_{l_i}}}{\epsilon_1 k \cos\theta_{l_i} + \epsilon_o\sqrt{k_1^2 - k^2 \sin^2\theta_{l_i}}} \tag{9}$$

Summing up the local parallel and perpendicular polarized components, we obtain

$$\hat{n} \times \overline{E}(\overline{r}') = E_o\Big\{(\hat{e}_i \cdot \hat{q}_i)(\hat{n} \times \hat{q}_i)(1 + R^{\mathrm{TE}}) + (\hat{e}_i \cdot \hat{p}_i)(\hat{n} \cdot \hat{k}_i)\hat{q}_i(1 - R^{\mathrm{TM}})\Big\} e^{i\overline{k}_i \cdot \overline{r}'} \tag{10a}$$

$$\hat{n} \times \overline{H}(\overline{r}') = \frac{E_o}{\eta}\Big\{-(\hat{e}_i \cdot \hat{q}_i)(\hat{n} \cdot \hat{k}_i)\hat{q}_i(1 - R^{\mathrm{TE}}) + (\hat{e}_i \cdot \hat{p}_i)(\hat{n} \times \hat{q}_i)(1 + R^{\mathrm{TM}})\Big\} e^{i\overline{k}_i \cdot \overline{r}'} \tag{10b}$$

The local angle of incidence can be calculated from the formula

$$\cos\theta_{l_i} = -\hat{n} \cdot \hat{k}_i \tag{11}$$

The normal vector at the point $\overline{r}'$ is given by

$$\hat{n}(\overline{r}') = \frac{-\hat{x}\alpha - \hat{y}\beta + \hat{z}}{\sqrt{1 + \alpha^2 + \beta^2}} \tag{12}$$

where α and β are the local slopes in the $\hat{x}$ and $\hat{y}$ directions,

$$\alpha = \frac{\partial f(x', y')}{\partial x'} \tag{13a}$$

$$\beta = \frac{\partial f(x', y')}{\partial y'} \tag{13b}$$

Substituting (10) into (5), we obtain, after some algebraic manipulations,

$$\overline{E}_s(\bar{r}) = \frac{ik\, e^{ikr}}{4\pi r} E_o(\overline{\overline{I}} - \hat{k}_s\hat{k}_s) \cdot \int_{A_o} d\bar{r}'_\perp\, \overline{F}(\alpha,\beta)\, e^{i(\bar{k}_i - \bar{k}_s)\cdot\bar{r}'} \tag{14a}$$

Similarly, for the transmitted field,

$$\overline{E}_t(\bar{r}) = -\frac{ik_1\, e^{ik_1 r}}{4\pi r} E_o(\overline{\overline{I}} - \hat{k}_t\hat{k}_t) \cdot \int_{A_o} d\bar{r}'_\perp\, \overline{N}(\alpha,\beta)\, e^{i(\bar{k}_i - \bar{k}_t)\cdot\bar{r}'} \tag{14b}$$

where

$$\begin{aligned}\overline{F}(\alpha,\beta) = (1+\alpha^2+\beta^2)^{1/2}\Big\{ & -(\hat{e}_i\cdot\hat{q}_i)(\hat{n}\cdot\hat{k}_i)\hat{q}_i(1-R^{\mathrm{TE}}) \\ & + (\hat{e}_i\cdot\hat{p}_i)(\hat{n}\times\hat{q}_i)(1+R^{\mathrm{TM}}) \\ & + (\hat{e}_i\cdot\hat{q}_i)(\hat{k}_s\times(\hat{n}\times\hat{q}_i))(1+R^{\mathrm{TE}}) \\ & + (\hat{e}_i\cdot\hat{p}_i)(\hat{n}\cdot\hat{k}_i)(\hat{k}_s\times\hat{q}_i)(1-R^{\mathrm{TM}})\Big\}\end{aligned} \tag{15a}$$

$$\begin{aligned}\overline{N}(\alpha,\beta) = (1+\alpha^2+\beta^2)^{1/2}\Big\{ & -\frac{\eta_1}{\eta}(\hat{e}_i\cdot\hat{q}_i)(\hat{n}\cdot\hat{k}_i)\hat{q}_i(1-R^{\mathrm{TE}}) \\ & + \frac{\eta_1}{\eta}(\hat{e}_i\cdot\hat{p}_i)(\hat{n}\times\hat{q}_i)(1+R^{\mathrm{TM}}) \\ & + (\hat{e}_i\cdot\hat{q}_i)[(\hat{k}_t\times(\hat{n}\times\hat{q}_i)](1+R^{\mathrm{TE}}) \\ & + (\hat{e}_i\cdot\hat{p}_i)(\hat{n}\cdot\hat{k}_i)(\hat{k}_t\times\hat{q}_i)(1-R^{\mathrm{TM}})\Big\}\end{aligned} \tag{15b}$$

We note that except for the phase factors, the expressions in the integrands of the diffraction integral (14) are not explicit functions of $\bar{r}'$. They are explicit functions of the slopes α and β which are functions of $\bar{r}'$. The tangent-plane-approximated diffraction integrals, as expressed in (14), do not take into account the effects of shadowing and multiple scattering.

Expanding the integrands $\overline{F}(\alpha,\beta)$ and $\overline{N}(\alpha,\beta)$ about zero slopes, we obtain

$$\overline{F}(\alpha,\beta) = \overline{F}(0,0) + \alpha \left.\frac{\partial\overline{F}}{\partial\alpha}\right|_{\alpha,\beta=0} + \beta \left.\frac{\partial\overline{F}}{\partial\beta}\right|_{\alpha,\beta=0} + \cdots \tag{16}$$

$$\overline{N}(\alpha,\beta) = \overline{N}(0,0) + \alpha \left.\frac{\partial \overline{N}}{\partial \alpha}\right|_{\alpha,\beta=0} + \beta \left.\frac{\partial \overline{N}}{\partial \beta}\right|_{\alpha,\beta=0} + \cdots \tag{17}$$

where $\overline{F}(0,0)$ and $\overline{N}(0,0)$ are evaluated at $\alpha = \beta = 0$, etc. For angles of incidence near normal and for surfaces with small root mean square (rms) slope, the Fresnel reflection coefficients vary only slightly with a change in local angle of incidence. Keeping only the first terms in (16) and (17), we obtain from (14)

$$\overline{E}_s = \frac{ik\, e^{ikr}}{4\pi r} E_o(\overline{\overline{I}} - \hat{k}_s\hat{k}_s) \cdot \overline{F}(0,0) I \tag{18a}$$

$$\overline{E}_t = -\frac{ik_1\, e^{ik_1 r}}{4\pi r} E_o(\overline{\overline{I}} - \hat{k}_t\hat{k}_t) \cdot \overline{N}(0,0) I_t \tag{18b}$$

where the integrals I and I_t are given by

$$I = \iint_{A_o} d\bar{r}'_\perp\, e^{i(\bar{k}_i - \bar{k}_s)\cdot\bar{r}'} \tag{19a}$$

$$I_t = \iint_{A_o} d\bar{r}'_\perp\, e^{i(\bar{k}_i - \bar{k}_t)\cdot\bar{r}'} \tag{19b}$$

The scattered and transmitted fields are next separated into a mean field and a fluctuating part of the field

$$\overline{E}_s(\bar{r}) = \overline{E}_{sm}(\bar{r}) + \overline{\mathcal{E}}_{sf}(\bar{r}) \tag{20a}$$

$$\overline{E}_t(\bar{r}) = \overline{E}_{tm}(\bar{r}) + \overline{\mathcal{E}}_{tf}(\bar{r}) \tag{20b}$$

with

$$<\overline{\mathcal{E}}_{sf}(\bar{r})> = <\overline{\mathcal{E}}_{tf}(\bar{r})> = 0$$

where $\overline{E}_{sm}$ and $\overline{E}_{tm}$ denote the mean scattered and transmitted fields respecively. The total scattered intensity is then a sum of coherent and incoherent intensities

$$<|\overline{E}_s(\bar{r})|^2> = |\overline{E}_{sm}|^2 + <|\overline{\mathcal{E}}_{sf}(\bar{r})|^2> \tag{21a}$$

$$<|\overline{E}_t(\bar{r})|^2> = |\overline{E}_{tm}|^2 + <|\overline{\mathcal{E}}_{tf}(\bar{r})|^2> \tag{21b}$$

In view of (18) and (19) and noting that $(\hat{p}_i, \hat{q}_i, \hat{k}_i)$ $(\hat{v}_s, \hat{h}_s, \hat{k}_s)$, and $(\hat{v}_t, \hat{h}_t, \hat{k}_t)$ are the three orthogonal systems for the incident, scattered,

and transmitted systems, we find

$$|\overline{E}_{sm}(\bar{r})|^2 = \frac{k^2|E_o|^2}{16\pi^2 r^2}\left\{\left|\hat{v}_s \cdot \overline{F}(0,0)\right|^2 + \left|\hat{h}_s \cdot \overline{F}(0,0)\right|^2\right\}|<I>|^2 \quad (22a)$$

$$<|\overline{\mathcal{E}}_{sf}(\bar{r})|^2> = \frac{k^2|E_o|^2}{16\pi^2 r^2}\left\{\left|\hat{v}_s \cdot \overline{F}(0,0)\right|^2 + \left|\hat{h}_s \cdot \overline{F}(0,0)\right|^2\right\} D_I \quad (22b)$$

$$|\overline{E}_{tm}(\bar{r})|^2 = \frac{k_1^2|E_o|^2}{16\pi^2 r^2}\left\{\left|\hat{v}_t \cdot \overline{N}(0,0)\right|^2 + \left|\hat{h}_t \cdot \overline{N}(0,0)\right|^2\right\}|<I_t>|^2 \quad (23a)$$

$$<|\overline{\mathcal{E}}_{tf}(\bar{r})|^2> = \frac{k_1^2|E_o|^2}{16\pi^2 r^2}\left\{\left|\hat{v}_t \cdot \overline{N}(0,0)\right|^2 + \left|\hat{h}_t \cdot \overline{N}(0,0)\right|^2\right\} D_{I_t} \quad (23b)$$

where

$$D_I = <|I|^2> - |<I>|^2 \quad (24a)$$

$$D_{I_t} = <|I_t|^2> - |<I_t>|^2 \quad (24b)$$

We now specify the height distribution function $f(\bar{r}_\perp)$ by assuming a stationary Gaussian process and that the probability for $f(\bar{r}_\perp)$ is independent of the position $\bar{r}_\perp$ on the rough surface and has the Gaussian distribution

$$p(f(\bar{r}_\perp)) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-f^2/2\sigma^2} \quad (25)$$

where σ is the standard deviation of the surface height. For two points $\bar{r}_{\perp 1}$ and $\bar{r}_{\perp 2}$ on the surface, the joint probability density [Davenport and Root, 1958] is

$$p(f_1(\bar{r}_{\perp 1}), f_2(\bar{r}_{\perp 2})) = \frac{e^{-(f_1^2 - 2Cf_1f_2 + f_2^2)/2\sigma^2(1-C^2)}}{2\pi\sigma^2\sqrt{1-C^2}} \quad (26)$$

where C is the correlation coefficient between the two points and is a function of $\bar{r}_{\perp 1}$ and $\bar{r}_{\perp 2}$. For a statistically homogeneous isotropic surface, C is only a function of $\rho = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$,

$$<f(\bar{r}_{\perp 1})f(\bar{r}_{\perp 2})> = \sigma^2 C(\rho) \quad (27)$$

with $C(0) = 1$ and $C(\infty) = 0$. It is easily shown that

$$< e^{i\nu f(\bar{r}_\perp)} > = \int_{-\infty}^{\infty} df\, p(f) e^{i\nu f} = e^{-\sigma^2\nu^2/2} \quad (28)$$

and

$$< e^{i\nu(f_1(\bar{r}_{\perp 1})-f_2(\bar{r}_{\perp 2}))} > = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} df_1\, df_2\; p(f_1, f_2)\, e^{i\nu(f_1-f_2)}$$
$$= e^{-\sigma^2\nu^2(1-C(\rho))} \tag{29}$$

The expressions for $|<I>|^2$, D_I, $|<I_t>|^2$ and D_{I_t} can now be derived in terms of the statistical moments of the height distribution.

The integral I is given by

$$I = \iint_{A_o} d\bar{r}'_\perp\, e^{i\bar{k}_{d\perp}\cdot\bar{r}'_\perp}\, e^{ik_{dz}f(\bar{r}'_\perp)} \tag{30}$$

where

$$\bar{k}_d = \bar{k}_i - \bar{k}_s = \hat{x}k_{dx} + \hat{y}k_{dy} + \hat{z}k_{dz} \tag{31}$$

The ensemble average of I is

$$<I> = \iint_{A_o} d\bar{r}'_\perp\, e^{i\bar{k}_{d\perp}\cdot\bar{r}'_\perp} < e^{ik_{dz}f(\bar{r}'_\perp)} >$$
$$= 4L_xL_y\, e^{-k_{dz}^2\sigma^2/2}\,\mathrm{sinc}(k_{dx}L_x)\,\mathrm{sinc}(k_{dy}L_y) \tag{32}$$

where $\mathrm{sinc}\, x = \sin x/x$, and $2L_x$ and $2L_y$ are the lengths of the rough surface illuminated in the $\hat{x}$ and $\hat{y}$ directions so that $A_o = 4L_xL_y$ By allowing L_x and L_y to approach infinity in the above expression, we obtain

$$|<I>|^2 = 4\pi^2 A_o e^{-k_{dz}^2\sigma^2}\delta(k_{dx})\delta(k_{dy}) \tag{33}$$

where we made use of the identity

$$\lim_{L_x, L_y\to\infty} \frac{L_x\, L_y}{\pi^2}\,\mathrm{sinc}(k_{dx}L_x)\,\mathrm{sinc}(k_{dy}L_y) = \delta(k_{dx})\,\delta(k_{dy})$$

The integral for $<II^*>$ is given by

$$<II^*> = \iint_{A_o} d\bar{r}_\perp \iint_{A_o} d\bar{r}'_\perp\; e^{i\bar{k}_{d\perp}\cdot(\bar{r}_\perp-\bar{r}'_\perp)} < e^{ik_{dz}(f(\bar{r}_\perp)-f(\bar{r}'_\perp))} >$$

Using (29) and making the usual change of variables to the difference and half the sum of coordinates, we obtain

$$<II^*> = \int_{-2L_x}^{2L_x} dx \int_{-2L_y}^{2L_y} dy\,(2L_x-|x|)(2L_y-|y|)$$
$$\cdot\, e^{ik_{dx}x+ik_{dy}y}\, e^{-k_{dz}^2\sigma^2(1-C(\rho))} \tag{34}$$

The correlation function $C(\rho)$ is assumed to have a Gaussian form

$$C(\rho) = e^{-\rho^2/l^2} \tag{35}$$

where l is the correlation length for the random variable $f(\bar{r}_\perp)$ in the transverse plane.

The expression for the standard derivation of the integral I can now be evaluated in closed form. First notice that $|<I>|^2$ can also be expressed as

$$|<I>|^2 = \int_{-2L_x}^{2L_x} dx \int_{-2L_y}^{2L_y} dy\,(2L_x - |x|)(2L_y - |y|)e^{ik_{dx}x + ik_{dy}y} e^{-\sigma^2 k_{dx}^2} \tag{36}$$

In view of (35), we note that the contribution of the integral of $< II^* > - |<I>|^2$ comes from $|x|$ and $|y|$ of the same order of l and the integrand is practically zero for $\rho = (x^2 + y^2)^{1/2}$ larger than a few l's. Assuming the illuminated rough surface contains many correlation lengths $L_x, L_y \gg l$, we obtain

$$\begin{aligned} D_I &= < II^* > - |<I>|^2 \\ &= A_o \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \left\{ e^{-\sigma^2 k_{dx}^2 (1 - C(\rho))} - e^{-\sigma^2 k_{dx}^2} \right\} e^{ik_{dx}x + ik_{dy}y} \end{aligned} \tag{37}$$

Converting the integral in (37) to cylindrical coordinates and carrying out the integral in $d\phi$ gives a Bessel function $J_o(k_\rho \rho)$ where $k_\rho = (k_{dx}^2 + k_{dy}^2)^{1/2}$ in the integrand. We further make a power series expansion

$$e^{-\sigma^2 k_{dx}^2 (1 - C(\rho))} - e^{-\sigma^2 k_{dx}^2} = e^{-\sigma^2 k_{dx}^2} \sum_{m=1}^{\infty} \frac{(\sigma^2 k_{dx}^2)^m}{m!} e^{-m\rho^2/l^2} \tag{38}$$

and make use of the integral identity

$$\int_o^{\infty} d\rho\, \rho\, J_o(k_\rho \rho)\, e^{-m\rho^2/l^2} = \frac{l^2}{2m} e^{-k_\rho^2 l^2/4m} \tag{39}$$

Using (38) and (39) in (37), we obtain

$$\begin{aligned} D_I &= < II^* > - |<I>|^2 \\ &= \pi A_o \sum_{m=1}^{\infty} \frac{(k_{dx}^2 \sigma^2)^m}{m! m} l^2\, e^{-(k_{dx}^2 + k_{dy}^2) l^2/4m}\, e^{-\sigma^2 k_{dx}^2} \end{aligned} \tag{40}$$

In a similar manner, the expressions for $|<I_t>|^2$ and D_{I_t} may be derived. They are

$$|<I_t>|^2 = 4\pi^2 A_o e^{-\sigma^2 k_{tdz}^2}\,\delta(k_{tdx})\delta(k_{tdy}) \tag{41}$$

and

$$D_{I_t} = \pi A_o \sum_{m=1}^{\infty} \frac{(k_{tdz}^2\sigma^2)^m}{m!m} l^2\, e^{-(k_{tdx}^2+k_{tdy}^2)l^2/4m}\, e^{-\sigma^2 k_{tdz}^2} \tag{42}$$

where

$$\overline{k}_{td} = \overline{k}_i - \overline{k}_t = \hat{x}k_{tdx} + \hat{y}k_{tdy} + \hat{z}k_{tdz}$$

The bistatic scattering coefficients for the reflected intensities are defined as

$$\gamma_{ab}^r(\hat{k}_s, \hat{k}_i) = \frac{4\pi r^2 (S_r)_a}{A_o \cos\theta_i (S_o)_b} \qquad (a,b = v,h) \tag{43}$$

where subscript b represents the polarization of the incident wave, subscript a the polarization of the scattered wave, S_o the Poynting power density of the incident wave, S_r the Poynting density of the scattered wave, A_o the area of the rough surface projected onto the $x-y$ plane, and θ_i the incident angle. From (18) and (22), we calculate the vertically and horizontally polarized coherent and incoherent scattered intensities for the cases of vertically and horizontally polarized incident fields. Let

$$\overline{F}_b(0,0) = \overline{F}(0,0)\Big|_{\hat{s}_i=\hat{b}_i}$$

$\overline{F}(0,0)$ can be calculated by setting $\alpha = \beta = 0$ in (12) and (15a). Next we take the dot product with $\hat{v}_s$ and $\hat{h}_s$. Thus,

$$\hat{h}_s \cdot \overline{F}_h(0,0) = \left[(1-R_o^{\mathrm{TE}})\cos\theta_i - (1+R_o^{\mathrm{TE}})\cos\theta_s\right]\cos(\phi_s-\phi_i) \tag{44a}$$

$$\hat{v}_s \cdot \overline{F}_h(0,0) = \left[(1-R_o^{\mathrm{TE}})\cos\theta_i\cos\theta_s - (1+R_o^{\mathrm{TE}})\right]\sin(\phi_s-\phi_i) \tag{44b}$$

$$\hat{h}_s \cdot \overline{F}_v(0,0) = \left[(1+R_o^{\mathrm{TM}}) - (1-R_o^{\mathrm{TM}})\cos\theta_i\cos\theta_s\right]\sin(\phi_s-\phi_i) \tag{44c}$$

$$\hat{v}_s \cdot \overline{F}_v(0,0) = \left[-(1+R_o^{\mathrm{TM}})\cos\theta_s + (1-R_o^{\mathrm{TM}})\cos\theta_i\right]\cos(\phi_s-\phi_i) \tag{44d}$$

where R_o^{TM} and R_o^{TE} are the Fresnel reflection coefficients of a smooth flat surface for the vertically and horizontally polarized incident waves.

In view of (21a), the bistatic scattering coefficients γ^r_{ab} can be decomposed into a coherent part and an incoherent part

$$\gamma^r_{ab}(\hat{k}_s, \hat{k}_i) = \frac{k^2}{4\pi A_o \cos\theta_i} \left|\hat{a}_s \cdot \overline{F}_b(0,0)\right|^2 \left\{ |<I>|^2 + D_I \right\} \tag{45}$$

The first term is the coherent part. Making use of (33) and (44) and the fact that

$$\delta(k_{dx})\delta(k_{dy}) = \frac{\delta(\theta_s - \theta_i)\delta(\phi_s - \phi_i)}{(k^2 \sin\theta_i \cos\theta_i)} \tag{46}$$

we find that the coherent part becomes

$$\frac{k^2}{4\pi A_o \cos\theta_i} \left|\hat{a}_s \cdot \overline{F}_b(0,0)\right|^2 |<I>|^2 = \frac{4\pi |R_{bo}|^2}{\sin\theta_i} e^{-4k^2\sigma^2 \cos^2\theta_i} \cdot \delta(\theta_s - \theta_i)\delta(\phi_s - \phi_i)\delta_{ab} \tag{47}$$

Thus the coherent wave exists only in the specular directions. A similar derivation and observation follows for the transmitted waves.

b. Geometrical Optics Solution

The diffraction integral in (14) can also be evaluated with the stationary-phase method which leads to the *geometrical optics* solution. The exponential phase factor in (14) is

$$\psi = \overline{k}_d \cdot \overline{r}' = k_{dx}x' + k_{dy}y' + k_{dz}f(x', y')$$

Setting $\partial\psi/\partial x' = 0$ and $\partial\psi/\partial y' = 0$, we find the stationary-phase points

$$\alpha_o = -\frac{k_{dx}}{k_{dz}}$$

$$\beta_o = -\frac{k_{dy}}{k_{dz}}$$

The slopes α_o and β_o are such that the incident and scattered wave directions form a specular reflection. This is seen from (12) which gives

$$\hat{n}(\alpha_o, \beta_o) = (\overline{k}_s - \overline{k}_i)/|\overline{k}_d|$$

Replacing the surface slopes α and β by α_o and β_o, we obtain from (14)

$$<|\overline{E}_s|^2> = \frac{k^2 |E_o|^2}{16\pi^2 r^2} \left|(\overline{\overline{I}} - \hat{k}_s\hat{k}_s) \cdot \overline{F}(\alpha_o, \beta_o)\right|^2 <II^*> \tag{48}$$

where

$$< II^* > = < \iint_{A_o} d\bar{r}_\perp \iint_{A_o} d\bar{r}'_\perp \, e^{i\bar{k}_{d\perp}\cdot(\bar{r}_\perp - \bar{r}'_\perp)} \, e^{ik_{dz}(f(\bar{r}_\perp) - f(\bar{r}'_\perp))} > \quad (49)$$

The above integral can be solved by the method of asymptotics. For large k, contributions of the integral come from regions where (x', y') is close to (x, y). Expanding $f(x', y')$ about (x, y),

$$f(x', y') = f(x, y) + \alpha(x' - x) + \beta(y' - y) + \cdots$$

and replacing the integration variables by

$$u = k(x - x')$$
$$v = k(y - y')$$

we obtain

$$< II^* > = \left\langle \frac{1}{k^2} A_o \iint du\, dv \, e^{iu(q_x + \alpha q_z) + iv(q_y + \beta q_z) + O(1/k)} \right\rangle$$

Ignoring the $O(1/k)$ and higher order terms, we have

$$< II^* > = \frac{4\pi^2 A_o}{k^2} < \delta(q_x + \alpha q_z)\delta(q_y + \beta q_z) >$$

Therefore

$$< \lim_{k\to\infty} II^* > = \frac{4\pi^2 A_o}{k^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\alpha\, d\beta \, \delta(q_x + \alpha q_z)\delta(q_y + \beta q_z)\, p(\alpha, \beta)$$

where $p(\alpha, \beta)$ is the probability density function for the slopes at the surface. It follows that

$$< \lim_{k\to\infty} II^* > = \frac{4\pi^2 A_o}{k_{dz}^2} p\left(-\frac{k_{dx}}{k_{dz}}, -\frac{k_{dy}}{k_{dz}}\right) \quad (50)$$

For the Gaussian random rough surface

$$p(\alpha, \beta) = \frac{1}{2\pi\sigma^2 |C''(0)|} \exp\left[-\frac{\alpha^2 + \beta^2}{2\sigma^2 |C''(0)|}\right] \quad (51)$$

where σ is the standard deviation of the height of rough surface and $C''(0)$ is the double derivative of the correlation function at $\rho = 0$. Thus, $\sigma^2|C''(0)|$ is the mean square surface slope s^2 and for the Gaussian correlation function of (35) with correlation length l,

$$s^2 = \sigma^2 |C''(0)| = 2\frac{\sigma^2}{l^2}$$

Using (51) in (50) gives

$$< II^* > = \frac{2\pi A_o}{k_{dz}^2 \sigma^2 |C''(0)|} e^{-(k_{dx}^2 + k_{dy}^2)/2k_{dz}^2\sigma^2|C''(0)|} \tag{52}$$

Another way to evaluate $< II^* >$ is to perform the ensemble average first, and then to approximate the integral. From (34)

$$< II^* > = \int_{-2L_x}^{2L_x} dx \int_{-2L_y}^{2L_y} dy\, (2L_x - |x|)(2L_y - |y|) e^{i\bar{k}_{d\perp}\cdot\bar{r}_\perp} e^{-k_{dz}^2\sigma^2(1-C(\rho))} \tag{53}$$

Since $k_{dz}^2\sigma^2 \gg 1$, most of the contribution comes from around the origin. Thus, expanding the integrand about the origin we have $1 - C(\rho) \approx \rho^2 |C''(0)|/2$, and substituting into (53) the integral can be evaluated readily by making use of the integral identity of (39). The final result for $< II^* >$ is the same as (52).

For an incident field with polarization b, the scattered intensity for polarization a_s is given by

$$< |E_s(\bar{r})|^2 > = \frac{k^2 |E_o|^2}{16\pi^2 r^2} |\hat{a}_s \cdot \overline{F}_b(\alpha_o, \beta_o)|^2 < II^* > \tag{54}$$

where

$$\overline{F}_b(\alpha_o, \beta_o) = \overline{F}(\alpha_o, \beta_o)\Big|_{\hat{e}_i = \hat{b}}$$

Using (15a), we find

$$|\hat{a}_s \cdot \overline{F}_b(\alpha_o, \beta_o)|^2 = \frac{|\bar{k}_d|^4}{k^2 |\hat{k}_i \times \hat{k}_s|^4 k_{dz}^2} f_{ba}$$

where

$$f_{vv} = |(\hat{h}_s \cdot \hat{k}_i)(\hat{h}_i \cdot \hat{k}_s) R^{\mathrm{TE}} + (\hat{v}_s \cdot \hat{k}_i)(\hat{v}_i \cdot \hat{k}_s) R^{\mathrm{TM}}|^2 \tag{55a}$$

$$f_{hv} = |(\hat{v}_s \cdot \hat{k}_i)(\hat{h}_i \cdot \hat{k}_s) R^{\mathrm{TE}} - (\hat{h}_s \cdot \hat{k}_i)(\hat{v}_i \cdot \hat{k}_s) R^{\mathrm{TM}}|^2 \tag{55b}$$

$$f_{vh} = |(\hat{h}_s \cdot \hat{k}_i)(\hat{v}_i \cdot \hat{k}_s) R^{\mathrm{TE}} - (\hat{v}_s \cdot \hat{k}_i)(\hat{h}_i \cdot \hat{k}_s) R^{\mathrm{TM}}|^2 \tag{55c}$$

$$f_{hh} = |(\hat{v}_s \cdot \hat{k}_i)(\hat{v}_i \cdot \hat{k}_s) R^{\mathrm{TE}} + (\hat{h}_s \cdot \hat{k}_i)(\hat{h}_i \cdot \hat{k}_s) R^{\mathrm{TM}}|^2 \tag{55d}$$

and R^{TM} and R^{TE} are evaluated at

$$\hat{n} = \frac{\hat{x}\, k_{dx}/k_{dz} + \hat{y}\, k_{dy}/k_{dz} + \hat{z}}{\left(k_{dx}^2/k_{dz}^2 + k_{dy}^2/k_{dz}^2 + 1\right)^{1/2}}$$

In view of (43) and (52) the bistatic scattering coefficients for the reflected intensities are

$$\gamma_{ab}^r(\hat{k}_s, \hat{k}_i) = \frac{f_{ab}|\overline{k}_d|^4}{\cos\theta_i |\hat{k}_i \times \hat{k}_s|^4 k_{dz}^4} \frac{e^{-(k_{dx}^2+k_{dy}^2)/2k_{dz}^2\sigma^2|C''(0)|}}{2\sigma^2\,|C''(0)|} \tag{56}$$

In the backscattering direction $\hat{k}_s = -\hat{k}_i$. The backscattering cross sections are defined to be

$$\sigma_{ab}(\hat{k}_i) = \cos\theta_i\ \gamma_{ab}^r(-\hat{k}_i, \hat{k}_i) \tag{57}$$

From (56), we obtain

$$\sigma_{hh}(\theta_i) = \sigma_{vv}(\theta_i) = \frac{|R|^2\, e^{-\tan^2\theta_i/2\sigma^2|C''(0)|}}{\cos^4\theta_i\, 2\sigma^2\,|C''(0)|} \tag{58}$$

$$\sigma_{vh}(\theta_i) = \sigma_{hv}(\theta_i) = 0 \tag{59}$$

where R is the reflection coefficient at normal incidence. We note that from (59) that there is no depolarization in the backscattering direction.

c. Small Perturbation Method

In the small perturbation method, use is made of the Huygens' principle in conjunction with the extinction theorem. We have

$$\iint_{S'} dS' \left\{ i\omega\mu_o \overline{\overline{G}}(\overline{r}, \overline{r}') \cdot [\hat{n} \times \overline{H}(\overline{r}')] + \nabla \times \overline{\overline{G}}(\overline{r}, \overline{r}') \cdot [\hat{n} \times \overline{E}(\overline{r}')] \right\}$$

$$+ \overline{E}_i(\overline{r}) = \begin{cases} \overline{E}(\overline{r}) & z > f(\overline{r}_\perp) \qquad (60a) \\ 0 & z < f(\overline{r}_\perp) \qquad (60b) \end{cases}$$

$$\iint_{S'} dS' \left\{ i\omega\mu_1 \overline{\overline{G}}_1(\overline{r}, \overline{r}') \cdot [\hat{n}_d \times \overline{H}_1(\overline{r}')] + \nabla \times \overline{\overline{G}}_1(\overline{r}, \overline{r}') \cdot [\hat{n}_d \times \overline{E}_1(\overline{r}')] \right\}$$

$$= \begin{cases} 0 & z > f(\overline{r}_\perp) \qquad (61a) \\ \overline{E}_1(\overline{r}) & z < f(\overline{r}_\perp) \qquad (61b) \end{cases}$$

Since tangential fields are continuous, we can define surface field unknowns

$$dS'\,\eta\,\hat{n}\times\overline{H}(\overline{r}')=d\overline{r}'_{\perp}\,\overline{a}(\overline{r}'_{\perp})=dS'\,\eta\,\hat{n}\times\overline{H}_1(\overline{r}') \tag{62a}$$

$$dS'\,\hat{n}\times\overline{E}(\overline{r}')=d\overline{r}'_{\perp}\,\overline{b}(\overline{r}'_{\perp})=dS'\,\hat{n}\times\overline{E}_1(\overline{r}') \tag{62b}$$

Next we make use of the integral representation of dyadic Green's function [Zuniga and Kong, 1980] (6.5.26)

$$\overline{\overline{G}}(\overline{r},\overline{r}')=-\hat{z}\hat{z}\frac{\delta(\overline{r},\overline{r}')}{k_o^2}$$
$$+\begin{cases}\dfrac{i}{8\pi^2}\displaystyle\iint d^2\overline{k}_{\perp}\frac{1}{k_z}[\hat{e}(k_z)\hat{e}(k_z)+\hat{h}(k_z)\hat{h}(k_z)]e^{i\overline{k}\cdot(\overline{r}-\overline{r}')} & z>z'\\ \dfrac{i}{8\pi^2}\displaystyle\iint d^2\overline{k}_{\perp}\frac{1}{k_z}[\hat{e}(-k_z)\hat{e}(-k_z)+\hat{h}(-k_z)\hat{h}(-k_z)]e^{i\overline{k}_1\cdot(\overline{r}-\overline{r}')} & z<z'\end{cases}$$

where $\hat{e}(-k_z)=\hat{e}(k_z)$ and $\hat{h}(-k_z)=\hat{e}\times\hat{k}_1/k$. Evaluating (60$b$) for $z<f_{min}$ and (61a) for $z>f_{max}$ we obtain

$$\overline{E}_i(\overline{r})=\frac{1}{8\pi^2}\int d\overline{k}_{\perp}\,e^{i\overline{k}_{\perp}\cdot\overline{r}_{\perp}}\,e^{-ik_zz}\frac{k}{k_z}\int d\overline{r}'_{\perp}\,e^{-i\overline{k}_{\perp}\cdot\overline{r}'_{\perp}}\,e^{ik_zf(\overline{r}'_{\perp})}$$
$$\cdot\Big\{\Big[\hat{e}(-k_z)\hat{e}(-k_z)+\hat{h}(-k_z)\hat{h}(-k_z)\Big]\cdot\overline{a}(\overline{r}'_{\perp})$$
$$+\Big[-\hat{h}(-k_z)\hat{e}(-k_z)+\hat{e}(-k_z)\hat{h}(-k_z)\Big]\cdot\overline{b}(\overline{r}'_{\perp})\Big\} \tag{63a}$$

$$0=\frac{1}{8\pi^2}\int d\overline{k}_{\perp}\,e^{i\overline{k}_{\perp}\cdot\overline{r}_{\perp}}\,e^{ik_{1z}z}\frac{k_1}{k_{1z}}\int d\overline{r}'_{\perp}\,e^{-i\overline{k}_{\perp}\cdot\overline{r}'_{\perp}}\,e^{-ik_{1z}f(\overline{r}'_{\perp})}$$
$$\cdot\Big\{\frac{k}{k_1}\Big[\hat{e}_1(k_{1z})\hat{e}_1(k_{1z})+\hat{h}_1(k_{1z})\hat{h}_1(k_{1z})\Big]\cdot\overline{a}(\overline{r}'_{\perp})$$
$$+\Big[-\hat{h}_1(k_{1z})\hat{e}_1(k_{1z})+\hat{e}_1(k_{1z})\hat{h}_1(k_{1z})\Big]\cdot\overline{b}(\overline{r}'_{\perp})\Big\} \tag{63b}$$

The above equations are the extended boundary conditions, and can be used to solve for the surface fields along with the following results of (62)

$$\hat{n}(\overline{r}'_{\perp})\cdot\overline{a}(\overline{r}'_{\perp})=0 \tag{64a}$$

$$\hat{n}(\bar{r}'_\perp) \cdot \bar{b}(\bar{r}'_\perp) = 0 \tag{64b}$$

Using (12), (64) can be rewritten as

$$a_z(\bar{r}'_\perp) = \left(\hat{x}\frac{\partial f(\bar{r}'_\perp)}{\partial x'} + \hat{y}\frac{\partial f(\bar{r}'_\perp)}{\partial y'}\right) \cdot \bar{a}_\perp(\bar{r}'_\perp) \tag{65a}$$

$$b_z(\bar{r}'_\perp) = \left(\hat{x}\frac{\partial f(\bar{r}'_\perp)}{\partial x'} + \hat{y}\frac{\partial f(\bar{r}'_\perp)}{\partial y'}\right) \cdot \bar{b}_\perp(\bar{r}'_\perp) \tag{65b}$$

with a_z and b_z as the z components of $\bar{a}$ and $\bar{b}$.

Once the surface fields are obtained, the scattered field in region 0 immediately follows from (60a)

$$\begin{aligned}\bar{E}_s(\bar{r}) = -\frac{1}{8\pi^2}\int d\bar{k}_\perp\, e^{i\bar{k}_\perp\cdot\bar{r}_\perp + ik_z z}\frac{k}{k_z}\int d\bar{r}'_\perp\, e^{-i\bar{k}_\perp\cdot\bar{r}'_\perp - ik_z f(\bar{r}'_\perp)} \\ \cdot\left\{\left[\hat{e}(k_z)\hat{e}(k_z) + \hat{h}(k_z)\hat{h}(k_z)\right]\cdot\bar{a}(\bar{r}'_\perp)\right. \\ \left. + \left[-\hat{h}(k_z)\hat{e}(k_z) + \hat{e}(k_z)\hat{h}(k_z)\right]\cdot\bar{b}(\bar{r}'_\perp)\right\}\end{aligned} \tag{66a}$$

To solve for the surface fields, the perturbation method makes use of series expansions. Let

$$\bar{a}(\bar{r}'_\perp) = \sum_{m=0}^{\infty}\frac{1}{m!}\bar{a}^{(m)}(\bar{r}'_\perp) \tag{67a}$$

$$\bar{b}(\bar{r}'_\perp) = \sum_{m=0}^{\infty}\frac{1}{m!}\bar{b}^{(m)}(\bar{r}'_\perp) \tag{67b}$$

where $\bar{a}^m$ and $\bar{b}^m$ are the m th-order solutions of $\bar{a}$ and $\bar{b}$. We also have

$$e^{\pm ik_z f(\bar{r}'_\perp)} = \sum_{m=0}^{\infty}\frac{1}{m!}[\pm ik_z f(\bar{r}'_\perp)]^m \tag{68a}$$

$$e^{\pm ik_{1z} f(\bar{r}'_\perp)} = \sum_{m=0}^{\infty}\frac{1}{m!}[\pm ik_{1z} f(\bar{r}'_\perp)]^m \tag{68b}$$

In SPM, f and its derivatives are regarded as small parameters. The expansion of (67) and (68) are substituted into (63) to obtain the set of equations for the different-order solutions. From (65) and (67)

$$a_z^{(0)}(\bar{r}'_\perp) = b_z^{(0)}(\bar{r}'_\perp) = 0 \tag{69a}$$

$$a_z^{(m)}(\bar{r}'_\perp) = m\left(\hat{x}\frac{\partial f(\bar{r}'_\perp)}{\partial x'} + \hat{y}\frac{\partial f(\bar{r}'_\perp)}{\partial y'}\right)\cdot \bar{a}_\perp^{(m-1)}(\bar{r}'_\perp) \tag{69b}$$

$$b_z^{(m)}(\bar{r}'_\perp) = m\left(\hat{x}\frac{\partial f(\bar{r}'_\perp)}{\partial x'} + \hat{y}\frac{\partial f(\bar{r}'_\perp)}{\partial y'}\right)\cdot \bar{b}_\perp^{(m-1)}(\bar{r}'_\perp) \tag{69c}$$

In summary, the assumptions for the SPM are

$$k_z f(\bar{r}'_\perp),\ k_{1z} f(\bar{r}'_\perp),\ \frac{\partial f}{\partial x'},\ \frac{\partial f}{\partial y'} \ll 1 \tag{70}$$

Substituting (67)–(68) into (63) and (65) and equating the same-order terms, we shall calculate the scattered fields to the zeroth and first-order in the following.

Zeroth-Order Solution

We first define an orthonormal system $(\hat{q}_i, \hat{p}_i, \hat{z}_i)$, such that

$$\hat{q}_i = \hat{x}\frac{k_{iy}}{k_{i\rho}} - \hat{y}\frac{k_{ix}}{k_{i\rho}} = \hat{e}(k_{iz}) \tag{71}$$

and $\hat{z}_i = \hat{z}$ and $\hat{p}_i = \hat{z}_i \times \hat{q}_i = (\hat{x}k_{ix} + \hat{y}k_{iy})/k_{\rho i}$ where $k_{\rho i}^2 = k_{ix}^2 + k_{iy}^2$ and let

$$\bar{a}(\bar{r}'_\perp) = \hat{q}_i a_q(\bar{r}'_\perp) + \hat{p}_i a_p(\bar{r}'_\perp) + \hat{z}_i a_z(\bar{r}'_\perp) \tag{72a}$$

$$\bar{b}(\bar{r}'_\perp) = \hat{q}_i b_q(\bar{r}'_\perp) + \hat{p}_i b_p(\bar{r}'_\perp) + \hat{z}_i b_z(\bar{r}'_\perp) \tag{72b}$$

and note that

$$\begin{aligned}\overline{E}_i(\bar{r}) &= \hat{e}_i\, E_o\, e^{i\bar{k}_{i\perp}\cdot\bar{r}_\perp - ik_{iz}z} \\ &= \frac{\hat{e}_i}{4\pi^2}\int d\bar{k}_\perp\, e^{i\bar{k}_\perp\cdot\bar{r}_\perp - ik_z z}\int d\bar{r}'_\perp\, e^{i\bar{k}_{i\perp}\cdot\bar{r}'_\perp - i\bar{k}_\perp\cdot\bar{r}'_\perp}\end{aligned} \tag{73}$$

Using (73) in (63a), we find

$$\hat{e}_i\, e^{i\bar{k}_{i\perp}\cdot\bar{r}'_\perp} = \frac{k}{2k_{iz}}\Big\{\Big[\hat{e}(-k_{iz})\hat{e}(-k_{iz}) + \hat{h}(-k_{iz})\hat{h}(-k_{iz})\Big]\cdot\bar{a}^{(0)}_\perp(\bar{r}'_\perp) + \Big[-\hat{h}(-k_{iz})\hat{e}(-k_{iz}) + \hat{e}(-k_{iz})\hat{h}(-k_{iz})\Big]\cdot\bar{b}^{(0)}_\perp(\bar{r}')\Big\} \tag{74a}$$

and from (63b), we have

$$\Big[\hat{e}_1(k_{1iz})\hat{e}_1(k_{1iz}) + \hat{h}_1(k_{1iz})\hat{h}_1(k_{1iz})\Big]\cdot\bar{a}^{(0)}_\perp(\bar{r}'_\perp)\frac{k}{k_1} + \Big[-\hat{h}_1(k_{1iz})\hat{e}_1(k_{1iz}) + \hat{e}_1(k_{1iz})\hat{h}_1(k_{1iz})\Big]\cdot\bar{b}^{(0)}_\perp(\bar{r}'_\perp) = 0 \tag{74b}$$

Using (69), (72), and noting that

$$\hat{e}(k_{iz}) = \frac{\hat{k}_i\times\hat{z}}{|\hat{k}_i\times\hat{z}|} = \frac{1}{k_{i\rho}}(\hat{x}k_{iy} - \hat{y}k_{ix})$$

$$\hat{h}(k_{iz}) = \frac{1}{k^2}\hat{e}\times\hat{k}_i = \frac{k_{iz}}{kk_{i\rho}}(\hat{x}k_{ix} + \hat{y}k_{iy}) + \hat{z}\frac{k_{i\rho}}{k}$$

we find from (74a)

$$\hat{e}_i\, e^{i\bar{k}_{i\perp}\cdot\bar{r}'} = \frac{k}{2k_{iz}}\Big\{\hat{e}(-k_{iz})\Big(a^{(0)}_q(\bar{r}') + \frac{k_{iz}}{k}b^{(0)}_p(\bar{r}'_\perp)\Big) + \hat{h}\ (-k_{iz})\Big(\frac{k_{iz}}{k}a^{(0)}_p(\bar{r}'_\perp) - b^{(0)}_q(\bar{r}'_\perp)\Big)\Big\} \tag{75a}$$

Using (74b), we have

$$ka^{(0)}_q(\bar{r}') - k_{1iz}b^{(0)}_p(\bar{r}') = 0 \tag{75b}$$

$$\frac{kk_{1iz}}{k_1^2}a^{(0)}_p(\bar{r}') + b^{(0)}_q(\bar{r}') = 0 \tag{75c}$$

Since (75a) contains two scalar equations, (75) provides four equations for the four unknowns $a^{(0)}_p$, $a^{(0)}_q$, $b^{(0)}_p$, and $b^{(0)}_q$. Solving them and substituting back into $\bar{a}^{(0)}(\bar{r}'_\perp)$ and $\bar{b}^{(0)}(\bar{r}'_\perp)$ gives

$$\bar{a}^{(0)}(\bar{r}'_\perp) = \bar{a}^{(0)}(\bar{k}_{i\perp})\, e^{i\bar{k}_{i\perp}\cdot\bar{r}'_\perp} \tag{76a}$$

$$\bar{b}^{(0)}(\bar{r}'_\perp) = \bar{b}^{(0)}(\bar{k}_{i\perp})\, e^{i\bar{k}_{i\perp}\cdot\bar{r}'_\perp} \tag{76b}$$

where

$$\bar{a}_q^{(0)}(\bar{k}_{i\perp}) = [\hat{e}(-k_{iz}) \cdot \hat{e}_i]\frac{k_{iz}}{k}(1 - R_o^{\mathrm{TE}}) \tag{77a}$$

$$\bar{a}_p^{(0)}(\bar{k}_{i\perp}) = [\hat{h}(-k_{iz}) \cdot \hat{e}_i](1 + R_o^{\mathrm{TM}}) \tag{77b}$$

$$\bar{b}_q^{(0)}(\bar{k}_{i\perp}) = -[\hat{h}(-k_{iz}) \cdot \hat{e}_i]\frac{k_{iz}}{k}(1 - R_o^{\mathrm{TM}}) \tag{77c}$$

$$\bar{b}_p^{(0)}(\bar{k}_{i\perp}) = [\hat{e}(-k_{iz}) \cdot \hat{e}_i](1 + R_o^{\mathrm{TE}}) \tag{77d}$$

and R_o^{TE} and R_o^{TM} are the Fresnel reflection coefficients for the TE and TM waves

$$R_o^{\mathrm{TE}} = \frac{k_{iz} - k_{1iz}}{k_{iz} + k_{1iz}} \tag{78a}$$

$$R_o^{\mathrm{TM}} = \frac{\epsilon_1 k_{iz} - \epsilon_o k_{1iz}}{\epsilon_1 k_{iz} + \epsilon_o k_{1iz}} \tag{78b}$$

Using (76) in (66), we find the zeroth-order scattered field to be

$$\begin{aligned}\bar{E}_s^{(0)} = \{&R_o^{\mathrm{TE}}[\hat{e}(-k_{iz}) \cdot \hat{e}_i]\hat{e}(k_{iz}) \\ &+ R_o^{\mathrm{TM}}[\hat{h}(-k_{iz}) \cdot \hat{e}_i]\hat{h}(k_{iz})\} E_o e^{i\bar{k}_{i\perp} \cdot \bar{r}_\perp + ik_{iz}z}\end{aligned} \tag{79}$$

which is the reflected field for a flat surface.

First-Order Solution

The first-order solution for the surface fields can be obtained by substituting (67) and (68) into (63), (65), (66), and (69) and equating first-order terms. From (69*a*) and (76*a*)

$$a_z^{(1)}(\bar{r}'_\perp) = \left(\hat{x}\frac{\partial f(\bar{r}'_\perp)}{\partial x'} + \hat{y}\frac{\partial f(\bar{r}'_\perp)}{\partial y'}\right) \cdot \bar{a}_\perp^{(0)}(\bar{k}_{i\perp})\, e^{i\bar{k}_{i\perp} \cdot \bar{r}'_\perp} \tag{80}$$

To simplify (80), we introduce the Fourier transforms

$$F(\bar{k}_\perp) = \frac{1}{(2\pi)^2}\int d\bar{r}'_\perp\, f(\bar{r}'_\perp)\, e^{-i\bar{k}_\perp \cdot \bar{r}'_\perp} \tag{81}$$

$$\bar{A}^{(1)}(\bar{k}_\perp) = \frac{1}{(2\pi)^2}\int d\bar{r}'_\perp\, \bar{a}^{(1)}(\bar{r}'_\perp)\, e^{-i\bar{k}_\perp \cdot \bar{r}'_\perp} \tag{82a}$$

$$\overline{B}^{(1)}(\overline{k}_\perp) = \frac{1}{(2\pi)^2}\int d\overline{r}'_\perp\, \overline{b}^{(1)}(\overline{r}'_\perp)\, e^{-i\overline{k}_\perp\cdot\overline{r}'_\perp} \tag{82b}$$

Strictly speaking, the Fourier transforms do not exist for random functions and stochastic Fourier Stieltjes integral have to be defined [Tatarskii, 1971; Ishimaru, 1978]. However, the final results for scattered intensities are not affected.

Multiply (80) by $e^{-i\overline{k}_\perp\cdot\overline{r}'_\perp}/(2\pi)^2$ and integrate over $d\overline{r}'_\perp$. We obtain, by expressing $\partial f(\overline{r}'_\perp)/\partial x'$ and $\partial f(\overline{r}'_\perp)/\partial y'$ in terms of $F(\overline{k}_\perp)$,

$$A_z^{(1)}(\overline{k}_\perp) = \left\{ \frac{k_x k_{iy} - k_y k_{ix}}{k_{i\rho}} a_q^{(0)}(\overline{k}_{i\perp}) + \left(\frac{k_x k_{ix} + k_y k_{iy}}{k_{i\rho}} - k_{i\rho} \right) a_p^{(0)}(\overline{k}_{i\perp}) \right\} iF(\overline{k}_\perp - \overline{k}_{i\perp}) \tag{83a}$$

Similarly from (69b)

$$B_z^{(1)}(\overline{k}_\perp) = \left\{ \frac{k_x k_{iy} - k_y k_{ix}}{k_{i\rho}} b_q^{(0)}(\overline{k}_{i\perp}) + \left(\frac{k_x k_{ix} + k_y k_{iy}}{k_{i\rho}} - k_{i\rho} \right) b_p^{(0)}(\overline{k}_{i\perp}) \right\} iF(\overline{k}_\perp - \overline{k}_{i\perp}) \tag{83b}$$

Next we match both sides of equation (63a) to the first order. We note that

$$\left[\int d\overline{r}'_\perp e^{-i\overline{k}_\perp\cdot\overline{r}'_\perp}\, e^{ik_z f(\overline{r}'_\perp)} \overline{a}(\overline{r}'_\perp) \right]_{\text{first order}}$$

$$= \int d\overline{r}'_\perp e^{-i\overline{k}_\perp\cdot\overline{r}'_\perp} \left[ik_z f(\overline{r}'_\perp)\overline{a}^{(0)}(\overline{k}_{i\perp}) e^{i\overline{k}_{i\perp}\cdot\overline{r}'} + \overline{a}^{(1)}(\overline{r}') \right]$$

$$= (2\pi)^2 \left[ik_z F(\overline{k}_\perp - \overline{k}_{i\perp})\overline{a}^{(0)}(\overline{k}_{i\perp}) + \overline{A}^{(1)}(\overline{k}_\perp) \right]$$

Hence, the first-order equation from (63a) is

$$\begin{aligned} 0 = {} & \left[\hat{e}(-k_z)\hat{e}(-k_z) + \hat{h}(-k_z)\hat{h}(-k_z) \right] \\ & \cdot \left[\overline{A}^{(1)}(\overline{k}_\perp) + ik_z \overline{a}^{(0)}(\overline{k}_{i\perp}) F(\overline{k}_\perp - \overline{k}_{i\perp}) \right] \\ & + \left[-\hat{h}(-k_z)\hat{e}(-k_z) + \hat{e}(-k_z)\hat{h}(-k_z) \right] \\ & \cdot \left[\overline{B}^{(1)}(\overline{k}_\perp) + ik_z \overline{b}^{(0)}(\overline{k}_{i\perp}) F(\overline{k}_\perp - \overline{k}_{i\perp}) \right] \end{aligned} \tag{84}$$

and similarly, from (63b) we obtain

$$0 = \frac{k}{k_1}\left[\hat{e}_1(k_{1z})\hat{e}_1(k_{1z}) + \hat{h}_1(k_{1z})\hat{h}_1(k_{1z})\right] \cdot \left[\overline{A}^{(1)}(\overline{k}_\perp) - ik_{1z}\overline{a}^{(0)}(\overline{k}_{i\perp})F(\overline{k}_\perp - \overline{k}_{i\perp})\right] + \left[-\hat{h}_1(k_{1z})\hat{e}_1(k_{1z}) + \hat{e}_1(k_{1z})\hat{h}_1(k_{1z})\right] \cdot \left[\overline{B}^{(1)}(\overline{k}_\perp) - ik_{1z}\overline{b}^{(0)}(\overline{k}_{i\perp})F(\overline{k}_\perp - \overline{k}_{i\perp})\right] \tag{85}$$

Equations (84) and (85) are vector equations so that they comprise four scalar equations for the four unknowns $A_q^{(1)}(\overline{k}_\perp)$, $A_p^{(1)}(\overline{k}_\perp)$, $B_q^{(1)}(\overline{k}_\perp)$, and $B_p^{(1)}(\overline{k}_\perp)$.

The first-order scattered fields can now be obtained from (66),

$$\overline{E}_s^{(1)} = -\frac{1}{2}\int d\overline{k}_\perp\, e^{i\overline{k}_\perp\cdot\overline{r}_\perp}\, e^{ik_z z}\,\frac{k}{k_z}\left\{\left[\hat{e}(k_z)\hat{e}(k_z) + \hat{h}(k_z)\hat{h}(k_z)\right] \cdot \left[\overline{A}^{(1)}(\overline{k}_\perp) - ik_z F(\overline{k}_\perp - \overline{k}_{i\perp})\overline{a}^{(0)}(\overline{k}_{i\perp})\right] + \left[-\hat{h}(k_z)\hat{e}(k_z) + \hat{e}(k_z)\hat{h}(k_z)\right] \cdot \left[\overline{B}^{(1)}(\overline{k}_\perp) - ik_z F(\overline{k}_\perp - \overline{k}_{i\perp})\overline{b}^{(0)}(\overline{k}_{i\perp})\right]\right\} \tag{86}$$

In view of the fact that

$$<\overline{F}(\overline{k}_\perp)> = \frac{1}{(2\pi)^2}\int d\overline{r}'_\perp\, e^{i\overline{k}_\perp\cdot\overline{r}'_\perp} <f(\overline{r}_\perp)> = 0$$

we find $<\overline{E}_s^{(1)}> = <\overline{E}_t^{(1)}> = 0$. Thus, the first-order solution does not modify the coherent reflection coefficient and we have to calculate the second-order solution to see the correction term for the coherent wave due to the rough surface.

The lowest-order incoherent coefficients can be derived from (86). For an incident field with polarization $\hat{a}_i$, the scattered intensity with polarization $\hat{b}_s$ is given by

$$<|E_s^{(1)}|^2> = \int d\overline{k}_\perp\, f'_{ba}\, W(|\overline{k}_\perp - \overline{k}_{i\perp}|) = \int d\Omega_s\, k^2 \cos\theta_s\, f'_{ba}\, W(|\overline{k}_\perp - \overline{k}_{i\perp}|) \tag{87}$$

where $W(|\bar{k}_\perp - \bar{k}_{i\perp}|)$ is the spectral density of the rough surface and is the Fourier transform of the correlation function.

The spectral density is

$$W(\bar{k}_\perp) = \frac{\sigma^2}{(2\pi)^2} \int d\bar{r}_\perp \, e^{i\bar{k}_\perp \cdot \bar{r}_\perp} C(\bar{r}_\perp)$$

and satisfies the relation

$$< F(\bar{k}'_\perp) F^*(\bar{k}_\perp) > = \delta(\bar{k}'_\perp - \bar{k}_\perp) W(|\bar{k}'_\perp|)$$

For a Gaussian correlation function of (35), we have the spectral density given by

$$W(|\bar{k}_\perp - \bar{k}_{\perp i}|) = \frac{1}{4\pi} \sigma^2 l^2 \, e^{-(k_{dx}^2 + k_{dy}^2) l^2/4} \tag{88}$$

where $\bar{k}_{d\perp} = \bar{k}_\perp - \bar{k}_{i\perp}$, σ is the standard deviation of the surface height and l is the correlation length for $f(\bar{r}_\perp)$ in the transverse plane.

The bistatic scattering coefficients $\gamma^r_{ba}(\hat{k}_s, \hat{k}_i)$ are defined as the ratio of the scattered power of polarization b_s per unit solid angle in the direction $\hat{k}_s$ to the intercepted power of polarization a_i in the direction $\hat{k}_i$ averaged over 4π radians. Thus, from (87) we have

$$\gamma^r_{ba}(\hat{k}_s, \hat{k}_i) = 4\pi \frac{k^2 \cos\theta_s f'_{ba} W(|\bar{k}_\perp - \bar{k}_{i\perp}|)}{\cos\theta_i |E_o|^2}$$

After carrying out a tedious solution for (84) and (85) and making use of (88), we obtain

$$\gamma^r_{ba}(\hat{k}_s, \hat{k}_i) = \frac{4k^4 \sigma^2 l^2 \cos^2\theta_s \cos^2\theta_i}{\cos\theta_i} f_{ba} e^{-k_{d\rho}^2 l^2/4} \tag{89}$$

where

$$k_{d\rho}^2 = k^2 \left[\sin^2\theta_s + \sin^2\theta_i - 2\sin\theta_s \sin\theta_i \cos(\phi_s - \phi_i) \right]$$

and

$$f_{hh} = \left| \frac{(k_1^2 - k^2)}{(k_z + k_{1z})(k_{iz} + k_{1iz})} \right|^2 \cos^2(\phi_s - \phi_i)$$

$$f_{vh} = \left| \frac{(k_1^2 - k^2)kk_{1z}}{(k_1^2 k_z + k^2 k_{1z})(k_{iz} + k_{1iz})} \right|^2 \sin^2(\phi_s - \phi_i)$$

$$f_{hv} = \left| \frac{(k_1^2 - k^2)kk_{1iz}}{(k_z + k_{1z})(k_1^2 k_{iz} + k^2 k_{1iz})} \right|^2 \sin^2(\phi_s - \phi_i)$$

$$f_{vv} = \left| \frac{(k_1^2 - k^2)}{(k_1^2 k_z + k^2 k_{1z})(k_1^2 k_{iz} + k^2 k_{1iz})} \right.$$
$$\left. \cdot \left[k_1^2 k^2 \sin\theta_s \sin\theta_i - k^2 k_{1z} k_{1iz} \cos(\phi_s - \phi_i) \right] \right|^2$$

In the backscattering direction $\hat{k}_s = -\hat{k}_i$. The backscattering cross sections per unit area are

$$\sigma_{hh} = 4k^4\sigma^2 l^2 \cos^4\theta_i \left| R_o^{\mathrm{TE}} \right|^2 e^{-k^2 l^2 \sin^2\theta_i} \tag{90a}$$

$$\sigma_{vv} = 4k^8\sigma^2 l^2 \cos^4\theta_i \left| \frac{(k_1^2 - k^2)(k_1^2 \sin^2\theta_i + k_{1z}k_{1zi})}{(k_1^2 k_{zi} + k^2 k_{1zi})^2} \right|^2 e^{-k^2 l^2 \sin^2\theta_i} \tag{90b}$$

$$\sigma_{vh} = \sigma_{hv} = 0 \tag{90c}$$

It is seen from (90c) that there is no depolarization in the backscattering direction. In order to calculate for the depolarization returns, we must resort to the solution of second-order fields.

6.7 Effective Permittivity for a Volume Scattering Medium

Electromagnetic waves are scattered upon entering a medium containing scatterers or inhomogeneities, which is referred to as a volume scattering medium. A homogeneous dielectric medium is characterized by a permittivity. For a volume scattering medium containing dielectric scatterers, we can account for its scattering effects by characterizing it with an effective permittivity.

Consider a homogeneous medium with permittivity ϵ containing small spheres of permittivity ϵ_s and radius a [Fig. 6.7.1]. Assume the electric field in the absence of the scatterers to be

$$\overline{E}_e = \hat{z} E_e = E_e(\hat{r}\cos\theta - \hat{\theta}\sin\theta) \tag{1}$$

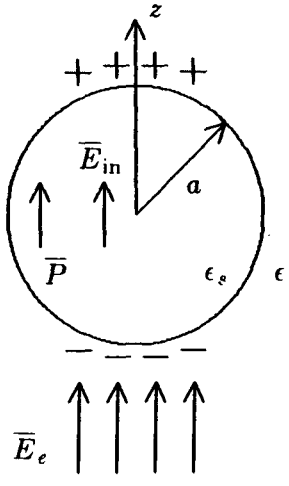


Figure 6.7.1 Sphere of radius a in a field $\overline{E}_e$.

In the static limit, the total field outside the sphere is

$$\overline{E}_{out} = \overline{E}_e + \frac{\epsilon_s - \epsilon}{\epsilon_s + 2\epsilon}\frac{a^3}{r^3}E_e(\hat{r}2\cos\theta + \hat{\theta}\sin\theta) \tag{2}$$

The total field inside the sphere is uniform and has the form

$$\overline{E}_{in} = \overline{E}_e - \frac{\epsilon_s - \epsilon}{\epsilon_s + 2\epsilon}\overline{E}_e = \frac{3\epsilon}{\epsilon_s + 2\epsilon}\overline{E}_e \tag{3}$$

It is easily shown that tangential $\overline{E}$ and normal $\overline{D}$ at $r = a$ are continuous.

The polarization vector for the sphere is

$$\overline{P} = (\epsilon_s - \epsilon)\overline{E}_{in} = 3\epsilon\frac{\epsilon_s - \epsilon}{\epsilon_s + 2\epsilon}\overline{E}_e \tag{4}$$

The polarizability of the sphere is seen to be

$$\alpha = 3\epsilon\frac{\epsilon_s - \epsilon}{\epsilon_s + 2\epsilon}\left(\frac{4\pi a^3}{3}\right) \tag{5}$$

The difference between the total fields inside and outside the sphere can be attributed to the polarization vector $\overline{P}$. From (3) and (4) we find

$$\overline{E}_e - \overline{E}_{in} = \frac{1}{3\epsilon}\overline{P} \tag{6}$$

Notice that $\overline{E}_e$ is the field in the absence of the sphere and constitutes only part of the total field outside the sphere as is evident from (2). We refer to $\overline{E}_e$ as the exciting field.

For a volume scattering medium containing n_0 spheres per unit volume, we relate the exciting field $\overline{E}_e$ and the macroscopic field $\overline{E}$ in the same manner as (6)

$$\overline{E}_e = \overline{E} + \frac{1}{3\epsilon}\overline{P} \tag{7}$$

The polarization vector is related to the exciting field by

$$\overline{P} = n_0 \alpha \overline{E}_e \tag{8}$$

In terms of the macroscopic field $\overline{E}$, we find from (7) and (8)

$$\overline{P} = \frac{n_0 \alpha}{1 - n_0 \alpha/3\epsilon}\overline{E} \tag{9}$$

The electric displacement vector $\overline{D}$ is

$$\overline{D} = \epsilon\overline{E} + \overline{P} = \frac{1 + 2n_0\alpha/3\epsilon}{1 - n_0\alpha/3\epsilon}\epsilon\overline{E}$$

It follows that the effective permittivity ϵ_{eff} is

$$\epsilon_{eff} = \epsilon\left[\frac{1 + 2n_0\alpha/3\epsilon}{1 - n_0\alpha/3\epsilon}\right] \tag{10}$$

The fractional volume occupied by the spheres is

$$f_s = n_0\left(\frac{4\pi a^3}{3}\right) \tag{11}$$

In view of (5), we write (9) in terms of the fractional volume f_s as follows:

$$\epsilon_{eff} = \epsilon\left[\frac{1 + 2f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}{1 - f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}\right] \tag{12}$$

Casting (10) in a more symmetric form, we obtain

$$\frac{\epsilon_{eff} - \epsilon}{\epsilon_{eff} + 2\epsilon} = f_s\frac{\epsilon_s - \epsilon}{\epsilon_s + 2\epsilon} \tag{13}$$

Equation (10) is known as the Clausius-Mossotti formula or the Lorentz-Lorenz formula. Equation (12) is known as the Maxwell-Garnett mixing formula and (13) as the Rayleigh mixing formula.

It is imporant to notice from (12) that as $f_s = 0$, $\epsilon_{eff} = \epsilon$ and as $f_s = 1$, $\epsilon_{eff} = \epsilon_s$. However, if both ϵ_s and ϵ are real, ϵ_{eff} is also real. This is because ϵ_{eff} as calculated above does not account for the scattering effect by the scatterers, which gives rise to an imaginary part for the effective permittivity.

a. Random Discrete Scatterers

In the low frequency limit, the scattered power can be attributed to the scattered power of the induced dipoles of the small spheres. From (5), (7), and (8), we find the relation between the excited field $\overline{E}_e$ and the macroscopic field $\overline{E}$ to be

$$\overline{E} = [1 - f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)]\,\overline{E}_e \tag{14}$$

Therefore the induced dipole moment for a small sphere is

$$\overline{p} = \alpha \overline{E}_e = \frac{4\pi a^3 \epsilon(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}{1 - f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)} \overline{E} \tag{15}$$

Assuming that the attenuation is small on the wavelength scale, we can approximate $\overline{E}$ by

$$\overline{E} = \hat{e} E_0 e^{i\overline{K}_R \cdot \overline{r}}$$

where K_R is the real part of the wavenumber $K = \omega(\mu_0 \epsilon_{eff})^{1/2}$. The induced dipole moment $\overline{p}_i$ of the i th scatterer centered at $\overline{r}_i$ is

$$\overline{p}_i = \hat{e} \frac{4\pi a^3 \epsilon(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}{1 - f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)} E_0 e^{i\overline{K}_R \cdot \overline{r}_i} \tag{16}$$

The scattered field of the i th dipole is

$$\overline{E}_{si}(\overline{r}) = \frac{\omega^2 \mu_0 e^{ikR_i}}{4\pi R_i} \left(\hat{R}_i \times \overline{p}_i\right) \times \hat{R}_i \tag{17}$$

where $\overline{R}_i = \overline{r} - \overline{r}_i$ is the vector pointing from the i th dipole to the observation point. For $r \gg r_i$, (17) can be approximated by

$$\overline{E}_{si}(\overline{r}) \approx \overline{A} e^{i(\overline{K}_R - \hat{r}k) \cdot \overline{r}_i}$$

where

$$\overline{A} = \frac{\omega^2 \mu_o e^{ikr}}{4\pi r}(\hat{r} \times \hat{e}) \times \hat{r} \frac{4\pi a^3 \epsilon(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}{1 - f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)} \tag{18}$$

The total scattered field $\overline{E}_s(\bar{r})$ is the sum of that due to all the dipoles,

$$\overline{E}_s(\bar{r}) = \sum_{i=1}^{N} \overline{A}\, e^{i(\overline{K}_R - \hat{r}k)\cdot \bar{r}_i} \tag{19}$$

where N is the total number of scatterers.

The total scattered intensity is determined to be

$$I_s = \frac{|\overline{A}|^2}{2\eta} \left\{ N + \sum_{i=1}^{N} \sum_{j \neq i} 2Re \left\{ e^{i\overline{K}\cdot(\bar{r}_i - \bar{r}_j)} \right\} \right\} \tag{20}$$

where $\overline{K} = \overline{K}_R - \hat{r}k$.

Notice that the position vectors $\bar{r}_i$ and $\bar{r}_j$ are random variables in space. Let $P_N(\bar{r}_1, \ldots, \bar{r}_N)$ be the N-particle probability density function for the scatterers center at $\bar{r}_1, \ldots, \bar{r}_N$. Taking the configuration average of (20) and assuming that all scatterers are identical, we find

$$< I_s > = \frac{|\overline{A}|^2}{2\eta} \{N + L\} \tag{21}$$

where

$$L = \sum_{i=1}^{N} \sum_{j \neq i} 2\, Re \left\{ \int d\bar{r}_1 \ldots \int d\bar{r}_N\, P_N(\bar{r}_1, \ldots, \bar{r}_N) e^{i\overline{K}\cdot(\bar{r}_i - \bar{r}_j)} \right\}$$

$$= N(N-1)\, Re \left\{ \int d\bar{r}_i \int d\bar{r}_j\, P_2(\bar{r}_i, \bar{r}_j) e^{i\overline{K}\cdot(\bar{r}_i - \bar{r}_j)} \right\}$$

The two-scatterer probability density function is, by Bayes' rule, $P_2(\bar{r}_i, \bar{r}_j) = P(\bar{r}_i | \bar{r}_j) P(\bar{r}_j)$ where $P(\bar{r}_i | \bar{r}_j)$ is the conditional probability of the i th scatterer at $\bar{r}_i$ given the j th scatterer at $\bar{r}_j$. We assume uniform distribution with $P(\bar{r}_j) = 1/V$ and $P(\bar{r}_i | \bar{r}_j) = g_2(\bar{r}_i, \bar{r}_j)/V$ where V is the volume containing the scatterers, and $g_2(\bar{r}_i, \bar{r}_j)$ is the two-point distribution function. For radially symmetric problems $g_2(\bar{r}_i, \bar{r}_j) = g(\bar{r}_i - \bar{r}_j)$ where g is called the pair-distribution function. We find

$$L = n_0 N\, Re \left\{ \int d\bar{r}\, g(\bar{r}) e^{i\overline{K}\cdot\bar{r}} \right\} \tag{22}$$

where $n_0 = N/V$.

The total scattered power is determined by integrating (21) over a 4π solid angle. Without loss of generality we let $\hat{e} = \hat{z}$ and obtain

$$P_s = \frac{E_0^2}{2\eta}\frac{8\pi}{3}k^4a^6\left|\frac{(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}{1 - f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}\right|^2 \{N + L\} \tag{23}$$

Consider a cylindrical volume of area A and length l containing n_0 scatterers per unit volume. The input power is $P_{\text{in}} = A(E_0^2/2)\sqrt{\epsilon_{eff}/\mu_o}$ $\approx AK_RE_0^2/2\eta k$ with ϵ_{eff} denoting the effective permittivity. The scattered power is given by (23) with $N = n_0Al$. The scattering induced attenuation rate is given by

$$2K_I = \frac{P_s}{lP_{\text{in}}} = 2f_s\frac{k^5a^3}{K_R}\left|\frac{(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}{1 - f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}\right|^2 \{1 + L/N\} \tag{24}$$

For $K_I \ll K_R$, we have $K^2 = \omega^2\mu_o(\epsilon_{effR} + i\epsilon_{effI}) \approx K_R^2 + i2K_RK_I$. The real part is obtained from (12) and the imaginary part from (24). We find the new complex effective permittivity to be

$$\epsilon_{eff} = \epsilon\left\{\left[\frac{1 + 2f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}{1 - f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}\right] + i2f_sk^3a^3\left|\frac{(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}{1 - f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}\right|^2 \{1 + L/N\}\right\} \tag{25}$$

The imaginary part of the complex wavenumber $K = K_R + iK_I$ now accounts for the scattering effects. As $K^2 = \omega^2\mu_o(\epsilon_{effR} + i\epsilon_{effI})$, the imaginary part of the complex effective permittivity ϵ_I is dependent on the pair distribution function $g(r)$ as seen from (22).

For independent scattering, the pair distribution function $g(\bar{r}) = 1$ and (22) can be converted into a surface integral at infinity bounding all the scatterers. As the wave solution satisfies the wave equation and the radiation condition, the result of the integral is zero,

$$\int d\bar{r}\, e^{i\bar{K}\cdot\bar{r}} = 0 \tag{26}$$

We find from (25), since $L = 0$,

$$\epsilon_{eff} = \epsilon\left\{\left[\frac{1 + 2f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}{1 - f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}\right] + i2f_sk^3a^3\left|\frac{(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}{1 - f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}\right|^2\right\} \tag{27}$$

The real part is the Maxwell-Garnett result and the imaginary part is identical to that obtained from the Rayleigh scattering result.

The hole-correction approximation is often used for impenetrable scatterers that are sparsely distributed (i.e., $f_s \ll 1$). The approximation requires that

$$\begin{aligned} g(\bar{r}) &= 0 \qquad \text{for} \qquad r < b \\ g(\bar{r}) &= 1 \qquad \text{for} \qquad r > b \end{aligned} \tag{28}$$

where $b = 2a$ is the distance separating the centers of two spheres, so that when the position of one sphere is given, the probability of another sphere being positioned within the distance b from its center is zero. But outside this exclusion volume of radius b, other spheres can be positioned anywhere else with equal probability as they are sparsely distributed. With this approximation, (22) becomes

$$L = n_0 N \, Re \left\{ \iiint\limits_{r>b} d\bar{r} \; e^{i\bar{K}\cdot\bar{r}} \right\}$$

The integral can be evaluated by converting it into a surface integral and neglecting the surface at infinity bounding the volume containing all the scatterers. We find

$$\begin{aligned} L &= n_0 N \, Re \left\{ \frac{1}{K^2} \iint d\bar{s} \cdot \nabla (e^{i\bar{K}\cdot\bar{r}}) \right\} \\ &= -\frac{n_0 N}{K} 2\pi b^2 \int_0^\pi d\theta \, \sin\theta \cos\theta \sin(Kb\cos\theta) \\ &= -n_0 N \frac{4\pi}{3} b^3 = -8 N f_s \end{aligned} \tag{29}$$

In obtaining the above result, we take the low frequency limit so that $Kb \ll 1$ and that $\sin(Kb\cos\theta) \approx Kb\cos\theta$.

The hole-correction result can also be obtained by making use of (26) in (22) which becomes

$$\begin{aligned} L &= n_0 N \, Re \iiint d\bar{r} \, [g(r) - 1] e^{i\bar{K}\cdot\bar{r}} \\ &\approx n_0 N \, Re \iiint d\bar{r} \, [g(r) - 1] \end{aligned} \tag{30}$$

where we again assumed at low frequencies, the exponential term is approximately equal to one since for large distance r, $g(r) \approx 1$. Making use of the hole-correction approximation in (28), we find $L = -n_0 N(4\pi b^3/3) = -8Nf_s$ which is the result in (29).

Substituting (29) into (25) we find the effective permittivity under the hole-correction (H-C) approximation to be

$$\epsilon_{eff} = \epsilon \left\{ \left[\frac{1 + 2f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}{1 - f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)} \right] + i2f_s k^3 a^3 \left| \frac{(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}{1 - f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)} \right|^2 (1 - 8f_s) \right\} \quad (31)$$

Clearly the H-C result is only valid for small fractional volumes. For $f > 1/8$, we observe from (31) that the imaginary part will be negative, a physically unacceptable result.

The Percus-Yevick pair distribution function [Wertheim, 1963] is a more realistic approximation which, when applied to (30), gives

$$L = N \left[\frac{(1 - f_s)^4}{(1 + 2f_s)^2} - 1 \right] \quad (32)$$

Substituting (32) into (25) yields

$$\epsilon_{eff} = \epsilon \left\{ \left[\frac{1 + 2f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}{1 - f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)} \right] + i2f_s k^3 a^3 \left| \frac{(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)}{1 - f_s(\epsilon_s - \epsilon)/(\epsilon_s + 2\epsilon)} \right|^2 \frac{(1 - f_s)^4}{(1 + 2f_s)^2} \right\} \quad (33)$$

Notice that (33) reduces to (31) for $f_s \ll 1$. Even more appealingly, we find from (33), that $\epsilon_{eff} = \epsilon$ as $f_s = 0$ and that $\epsilon_{eff} = \epsilon_s$ as $f_s = 1$.

b. Effective Permittivity for a Continuous Random Medium

Consider a random medium with permittivity $\epsilon(\bar{r})$. Introducing an auxiliary permittivity ϵ_g, we write the vector wave equation as

$$\nabla \times \nabla \times \overline{E} - k_g^2 \overline{E} = k_o^2 \left(\frac{\epsilon(\bar{r}) - \epsilon_g}{\epsilon_o} \right) \overline{E} \quad (34)$$

where $k_g^2 = \omega^2 \mu_o \epsilon_g$ and $k_o^2 = \omega^2 \mu_o \epsilon_o$. Let $\overline{\overline{G}}_g(\bar{r},\bar{r}')$ be the dyadic Green's function that satisfies

$$\nabla \times \nabla \times \overline{\overline{G}}_g(\bar{r},\bar{r}') - k_g^2 \overline{\overline{G}}_g(\bar{r},\bar{r}') = \overline{\overline{I}}\delta(\bar{r}-\bar{r}') \tag{35}$$

Let $\overline{E}_0(\bar{r})$ denote the field solution for the homogeneous wave equation with wavenumber k_g in (34). We have

$$\overline{E}(\bar{r}) = \overline{E}_0(\bar{r}) + k_o^2 \iiint d^3\bar{r}' \, \overline{\overline{G}}_g(\bar{r},\bar{r}') \frac{\epsilon(\bar{r}') - \epsilon_g}{\epsilon_o} \cdot \overline{E}(\bar{r}') \tag{36}$$

Notice that the field point $\bar{r}$ and the source point $\bar{r}'$ may coincide and we must account for the singularities of the dyadic Green's functions.

The singularity of the dyadic Green's function depends on the shape of the infinitesimal exclusion volume. We assume the correlation function for the random medium to be spherically symmetric and choose a spherically-shaped exclusion volume. The dyadic Green's function can be decomposed into [Tsang et al., 1985]

$$\overline{\overline{G}}_g(\bar{r},\bar{r}') = \mathrm{PV}\overline{\overline{G}}_g(\bar{r},\bar{r}') - \overline{\overline{I}}\frac{1}{3k_g^2}\delta(\bar{r}-\bar{r}') \tag{37}$$

where PV stands for principal value. Substituting (37) into (36), we obtain

$$\overline{E}_e(\bar{r}) = \overline{E}_0(\bar{r}) + k_o^2 \iiint d^3\bar{r}' \, \mathrm{PV}\overline{\overline{G}}_g(\bar{r},\bar{r}') \, \xi(\bar{r}')\overline{E}_e(\bar{r}') \tag{38}$$

where

$$\overline{E}_e(\bar{r}) = \frac{\epsilon(\bar{r}) + 2\epsilon_g}{3\epsilon_g} \overline{E}(\bar{r}) \tag{39}$$

and

$$\xi(\bar{r}) = 3\frac{\epsilon_g}{\epsilon_o}\left[\frac{\epsilon(\bar{r}) - \epsilon_g}{\epsilon(\bar{r}) + 2\epsilon_g}\right] \tag{40}$$

Equation (39) represents the relation for the uniform electric field $\overline{E}$ inside a dielectric sphere when it is placed in an external static field $\overline{E}_e$. From (39), we find

$$\overline{E}_e(\bar{r}) = \frac{1}{3\epsilon_g}\overline{D}(\bar{r}) + \frac{2}{3}\overline{E}(\bar{r}) \tag{41}$$

where $\overline{D}(\bar{r}) = \epsilon(\bar{r})\overline{E}(\bar{r})$. Multiplication of (39) and (40) yields

$$\overline{D}(\bar{r}) = \epsilon_g \overline{E}(\bar{r}) + \epsilon_o \xi(\bar{r})\overline{E}_e(\bar{r}) \tag{42}$$

Applying ensemble averaging to the above equation gives

$$< \overline{D}(\bar{r}) > = \epsilon_g < \overline{E}(\bar{r}) > + \epsilon_o < \xi(\bar{r})\overline{E}_e(\bar{r}') > \tag{43}$$

Thus $< \xi(\bar{r})\overline{E}_e(\bar{r}) >$ plays the role of a polarization vector in a background medium of ϵ_g.

In the strong fluctuation theory, we impose the condition

$$< \xi(\bar{r}) > = 0 \tag{44}$$

By virtue of (40), we require

$$< \frac{\epsilon(\bar{r}) - \epsilon_g}{\epsilon(\bar{r}) + 2\epsilon_g} > = 0 \tag{45}$$

Suppose there are n constituents in a mixture constituting the random medium with permittivity ϵ_p and fractional volume f_p, where $p = 1, 2, \ldots, n$. We have

$$\sum_{p=1}^{n} f_p = 1 \tag{46}$$

Equation (45) gives

$$\sum_{p=1}^{n} \frac{\epsilon_p - \epsilon_g}{\epsilon_p + 2\epsilon_g} f_p = 0 \tag{47}$$

Algebraic manipulation of (46) and (47) leads directly to the Polder-van Santen mixing formula

$$\sum_{p=1}^{n} \frac{\epsilon_p - \epsilon_o}{\epsilon_p + 2\epsilon_g} f_p = \frac{\epsilon_g - \epsilon_o}{3\epsilon_g} \tag{48}$$

which was orginally derived [Polder and van Santen, 1946] by regarding the inhomogeneities as dipoles. We have thus established that ϵ_g is the effective permittivity in the low-frequency limit.

As frequency increases, the multiple scattering gives rise to an imaginary part in the effective permittivity to account for the wave attenuation. Taking ensemble average of (38) we obtain

$$< \overline{E}_e(\bar{r}) > = \overline{E}_0(\bar{r}) + k_o^2 \iiint d^3\bar{r}' \ \mathrm{PV}\overline{\overline{G}}_g(\bar{r},\bar{r}')\cdot < \xi(\bar{r}')\overline{E}_e(\bar{r}') > \quad (49)$$

The bilocal approximation applied to (38) yields

$$< \overline{E}_e(\bar{r}) > \approx \overline{E}_0(\bar{r}) + k_o^2 \iiint d^3\bar{r}' \iiint d^3\bar{r}'' \ \mathrm{PV}\overline{\overline{G}}_g(\bar{r},\bar{r}') \cdot \overline{\overline{\xi}}_{eff}(\bar{r}' - \bar{r}'')\cdot < \overline{E}_e(\bar{r}'') > \quad (50)$$

where

$$\overline{\overline{\xi}}_{eff}(\bar{r}' - \bar{r}'') = k_o^2 \ \mathrm{PV}\overline{\overline{G}}_g(\bar{r}',\bar{r}'')R_\xi(|\bar{r}' - \bar{r}''|) \quad (51)$$
$$R_\xi(|\bar{r}' - \bar{r}''|) = < \xi(\bar{r}')\xi(\bar{r}'') >$$

Comparison of (49) and (50) gives

$$< \xi(\bar{r}')\overline{E}_e(\bar{r}') > = \iiint d^3\bar{r}'' \ \overline{\overline{\xi}}_{eff}(\bar{r}' - \bar{r}'')\cdot < \overline{E}_e(\bar{r}'') > \quad (52)$$

We find from (41)

$$< \overline{E}_e(\bar{r}) > = \frac{1}{3\epsilon_g} < \overline{D}(\bar{r}) > + \frac{2}{3} < \overline{E}(\bar{r}) > \quad (53)$$

Analogous to (52), we define the effective permittivity $\overline{\overline{\epsilon}}_{eff}$ such that

$$< \overline{D}(\bar{r}) > = \iiint d^3\bar{r}' \ \overline{\overline{\epsilon}}_{eff}(\bar{r} - \bar{r}')\cdot < \overline{E}(\bar{r}') > \quad (54)$$

Substituting (53) into (52) and (52) into (43), we obtain

$$< \overline{D}(\bar{r}) > = \epsilon_g < \overline{E}(\bar{r}) > + \frac{\epsilon_o}{3\epsilon_g} \iiint d^3\bar{r}' \ \overline{\overline{\xi}}_{eff}(\bar{r} - \bar{r}')\cdot < \overline{D}(\bar{r}') > + \frac{2\epsilon_o}{3} \iiint d^3\bar{r}' \ \overline{\overline{\xi}}_{eff}(\bar{r} - \bar{r}')\cdot < \overline{E}(\bar{r}') > \quad (55)$$

Fourier transforming $\overline{D}$, $\overline{\overline{\epsilon}}_{eff}$, $\overline{E}$, $\overline{\overline{\xi}}_{eff}$, and $\overline{E}_e$, e.g.,

$$\overline{D}(\overline{k}) = \iiint d^3\overline{r}\ \overline{D}(\overline{r})e^{-i\overline{k}\cdot r}$$

we find from (54) and (55),

$$\left[\overline{\overline{I}} - \frac{\epsilon_o}{3\epsilon_g}\overline{\overline{\xi}}_{eff}(\overline{k})\right] < \overline{D}(\overline{k}) > = \epsilon_g < \overline{E}(\overline{k}) > + \frac{2\epsilon_o}{3}\overline{\overline{\xi}}_{eff}(\overline{k}) < \overline{E}(\overline{k}) > \tag{56}$$

Thus

$$\overline{\overline{\epsilon}}_{eff}(\overline{k}) = \epsilon_g\overline{\overline{I}} + \epsilon_o\left[\overline{\overline{I}} - \frac{\epsilon_o}{3\epsilon_g}\overline{\overline{\xi}}_{eff}(\overline{k})\right]^{-1}\overline{\overline{\xi}}_{eff}(\overline{k}) \tag{57}$$

The second term is the correction provided by the bilocal approximation the validity of which is substantiated by

$$|\overline{\overline{\xi}}_{eff}(\overline{k})| \ll 1 \tag{58}$$

We obtain from (57)

$$\overline{\overline{\epsilon}}_{eff} \approx \epsilon_g\overline{\overline{I}} + \epsilon_o\overline{\overline{\xi}}^{(0)}_{eff} \tag{59}$$

where

$$\overline{\overline{\xi}}^{(0)}_{eff} = k_o^2 \iiint d^3\overline{r}\ \mathrm{PV}\overline{\overline{G}}_g(\overline{r})\, R_\xi(|\overline{r}|)$$

follows from the Fourier transform of (51) in which we let $e^{ikr} \approx 1$, assuming low frequency and small r. For spherical scatterers, we calculate

$$\begin{aligned}\mathrm{PV}\overline{\overline{G}}_g(\overline{r}) &= (\overline{\overline{I}} + \frac{1}{k_g^2}\nabla\nabla)\frac{e^{ik_gr}}{r} \\ &= \overline{\overline{I}}(-1 + ik_gr + k_g^2r^2)\frac{e^{ik_gr}}{4\pi k_g^2r^3} \\ &\quad + \hat{r}\hat{r}(3 - i3k_gr - k_g^2r^2)\frac{e^{ik_gr}}{4\pi k_g^2r^3}\end{aligned}$$

where $\hat{r} = \hat{x}x + \hat{y}y + \hat{z}z$. We find

$$\begin{aligned}
\bar{\bar{\xi}}^{(0)}_{eff} &= \bar{\bar{I}}k_o^2 \int dr\,(-1 + ik_g r + k_g^2 r^2)\frac{e^{ik_g r}}{k_g^2 r} R_\xi(|\bar{r}|) \\
&\quad + k_o^2 \int\int_0^\pi \int_0^{2\pi} r^2\, dr d\theta d\phi \sin\theta (3 - i3k_g r - k_g^2 r^2)\frac{e^{ik_g r}}{4\pi k_g^2 r^3} R_\xi(|\bar{r}|)\hat{r}\hat{r} \\
&= \bar{\bar{I}}k_o^2 \int dr\,(-1 + ik_g r + k_g^2 r^2)\frac{e^{ik_g r}}{k_g^2 r} R_\xi(|\bar{r}|) \\
&\quad + \bar{\bar{I}}k_o^2 \int dr\,(3 - i3k_g r - k_g^2 r^2)\frac{e^{ik_g r}}{3k_g^2 r} R_\xi(|\bar{r}|) \\
&= \bar{\bar{I}}\frac{2}{3}k_o^2 \int dr\, r\, R_\xi(|\bar{r}|)\, e^{ik_g r}
\end{aligned}$$

In the low-frequency limit we approximate $e^{ik_g r} \approx 1 + ik_g r$. Thus we obtain

$$\epsilon_{eff} \approx \epsilon_g + i\frac{2\epsilon_o}{3}k_o^2 k_g \int_0^a dr\, r^2 R_\xi(r)$$

For two species with fractional volumes f and f_s for permittivities ϵ and ϵ_s, respectively,

$$R_\xi(r) = <\xi^2> = f_s\left[3\frac{\epsilon_g}{\epsilon}\left(\frac{\epsilon_s - \epsilon_g}{\epsilon_s + 2\epsilon_g}\right)\right]^2 + (1 - f_s)\left[3\frac{\epsilon_g}{\epsilon}\left(\frac{\epsilon - \epsilon_g}{\epsilon + 2\epsilon_g}\right)\right]^2$$

Therefore, in the low-frequency limit, we obtain the effective permittity

$$\epsilon_{eff} = \epsilon_g\left\{1 + i2k_g^3 a^3\left[f_s\left(\frac{\epsilon_s - \epsilon_g}{\epsilon_s + 2\epsilon_g}\right)^2 + (1 - f_s)\left(\frac{\epsilon - \epsilon_g}{\epsilon + 2\epsilon_g}\right)^2\right]\right\} \tag{60}$$

Remarkably this formula also yields the limiting values of $\epsilon_{eff} = \epsilon_s$ as $f_s = 1$ and $\epsilon_{eff} = \epsilon$ as $f_s = 0$.

For small fractional volume, we let $f_s \ll 1$ and $\epsilon_g \approx \epsilon + \delta$, with δ denoting a small number. The above result reduces to

$$\begin{aligned}
\epsilon_g &\approx \epsilon[1 + 3f_s(\epsilon_s - \epsilon_g)/(\epsilon_s + 2\epsilon_g)] \\
\epsilon_{eff} &\approx \epsilon\left\{1 + 3f_s\left(\frac{\epsilon_s - \epsilon_g}{\epsilon_s + 2\epsilon_g}\right) + i2f_s k^3 a^3\left(\frac{\epsilon_s - \epsilon_g}{\epsilon_s + 2\epsilon_g}\right)^2\right\}
\end{aligned}$$

This is the same as that obtained with the random discrete scatterer model.

PROBLEMS

Problem P6.1

Consider a current distribution which is uniform within a cylinder of radius b and length $2a$ [Fig. P6.1] and flows in the $\hat{z}$ direction (such a current distribution might be thought of as an electron beam). In order to find the electric field at the origin i.e., at the center of the current distribution, we make use of the vector and scalar potentials $\overline{A}$ and ϕ. First we surround the origin O by a small cavity ΔV. The vector potential $\overline{A}$, which satisfies the homogeneous Helmholtz equations, is obtained from

$$\overline{A}(\overline{r}_0) = \frac{\mu}{4\pi} \iiint_{V-\Delta V} dV\, \overline{J}(\overline{r}) \frac{e^{ik|\overline{r}-\overline{r}_0|}}{|\overline{r}-\overline{r}_0|}$$

The scalar potential is obtained from

$$\phi(\overline{r}_0) = \frac{1}{4\pi\epsilon} \left[\iiint_{V-\Delta V} dV \rho(\overline{r}) \frac{e^{ik|\overline{r}-\overline{r}_0|}}{|\overline{r}-\overline{r}_0|} + \iint_S dS\, \frac{J(\overline{r})}{i\omega} \frac{e^{ik|\overline{r}-\overline{r}_0|}}{|\overline{r}-\overline{r}_0|} \right]$$

where $\rho(\overline{r}) = \nabla \cdot \overline{J}/i\omega = 0$ and $d\hat{S} \cdot \overline{J} = J$ at $z = \pm a$.

(a) Show that as $\Delta V \to 0$, the electric field at the origin is

$$\overline{E}(0) = \hat{z} J \left[\frac{i}{\omega\epsilon}(e^{ika} - 1) + \sqrt{\frac{\mu}{\epsilon}} \int_0^a dz\, e^{ik\sqrt{z^2+b^2}} \right] + \hat{z} \frac{iJa}{\omega\epsilon} \left[-\frac{e^{ika}}{a} + \frac{e^{ik\sqrt{a^2+b^2}}}{\sqrt{a^2+b^2}} \right]$$

where the first term is due to the vector potential $\overline{A}$ and the second term due to the scalar potential ϕ. Notice that the above result is independent of the shape of ΔV. Also evaluate $\overline{A}(\overline{r}_0)$ for any $\overline{r}_0$ inside the cylinder and then find $\overline{E}$ in terms of $\overline{A}(\overline{r}_0)$ by using

$$\overline{E} = i\omega \left(\overline{\overline{I}} + \frac{1}{k^2} \nabla\nabla \right) \cdot \overline{A}$$

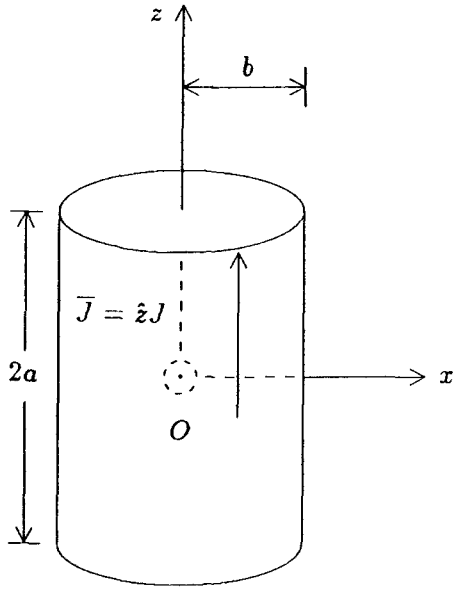


Figure P6.1

(b) To make use of the dyadic Green's function formalism to solve this problem, we first notice that

$$\begin{aligned}\bar{E}(\bar{r}) &= i\omega\mu \iiint_V dV' \, \bar{\bar{G}}(\bar{r},\bar{r}') \cdot \bar{J}(\bar{r}') \\ &= i\omega\mu \iiint_V dV' \left[\bar{\bar{I}} + \frac{1}{k^2}\nabla\nabla\right] \cdot \frac{\hat{z}Je^{ik|\bar{r}-\bar{r}'|}}{4\pi|\bar{r}-\bar{r}'|}\end{aligned}$$

Show that the first term is identical to that evaluated in (a).

(c) The second term in the integrand does not converge, because of the derivatives on $|\bar{r}-\bar{r}'|$ in the denominator. The integral is evaluated by excluding a small volume ΔV surrounding the field point to obtain its principal value and then letting $\Delta V \to 0$. The result depends critically on the shape of ΔV. To illustrate, consider ΔV to be a sphere of radius δ. Show that the second term gives

$$\bar{E} = \hat{z}\frac{iJa}{\omega\epsilon}\left[-\frac{e^{ika}}{a} + \frac{e^{ik\sqrt{a^2+b^2}}}{\sqrt{a^2+b^2}} + \frac{1}{3a}\right]$$

Notice the additional term of $1/3a$ in the above expression as compared with the second term in (a).

Thus, to make use of dyadic Green's functions when the observation point is in the source region, extreme caution must be exercised. We write for $\bar{r}_0$ in the source region,

$$\overline{E}(\bar{r}_0) = \text{PV} \iiint_V dV\, \overline{\overline{G}}(\bar{r}_0, \bar{r}') \cdot \overline{J}(\bar{r}') + \overline{E}_C(\bar{r}_0)$$

where PV denotes the principal value and $\overline{E}_C(\bar{r}_0)$ the correction term, which depends on the shape of ΔV. The final result, which is the sum of the PV integral and $\overline{E}_C(\bar{r}_0)$, however, is independent of the shape of ΔV.

Problem P6.2

The electromagnetic field of a Hertzian dipole with current moment $Il = -i\omega p$ has been calculated.

(a) Determine the field of a static dipole with moment p by letting $\omega \to 0$ in the above expressions. Show that $\overline{H} = 0$ and

$$\overline{E} = \frac{p}{4\pi\epsilon r^3}\left(\hat{\theta}\sin\theta + \hat{r}\,2\cos\theta\right)$$

Sketch the static field pattern.

(b) Consider the Rayleigh scattering of electromagnetic waves by particles of size much smaller than a wavelength, such as sunlight by air molecules. Model the particle as a small sphere of radius a and permittivity ϵ_p. Show that when the particle is illuminated with a light wave with electric field intensity E_0 in the $\hat{z}$ direction, the induced dipole moment p is

$$p = 4\pi\epsilon_o a^3 \left(\frac{\epsilon_p - \epsilon_o}{\epsilon_p + 2\epsilon_o}\right) E_0$$

(c) Show that the total power re-radiated by the particle acting as a Hertzian dipole is

$$P = \frac{4\pi}{3\eta}\left[\frac{\epsilon_p - \epsilon_o}{\epsilon_p + 2\epsilon_o}\right]^2 k^4 a^6 E_0^2$$

Find the scattering cross section defined by $2\eta P/E_0^2$. The above result is usually used to explain why the sky is blue (but why isn't it purple?).

(d) Consider an optical fiber with cross section radius R. The electromagnetic wave guided inside the fiber is scattered by the atoms and the molecules making up the fiber. Since the size of the scattering particles are much smaller than the guided light wavelength, the process can again be described by Rayleigh scattering. Assume the guided light has intensity E_0, wavelength $10^{-6}\,\mathrm{m}$, particle radius $a = 10^{-10}\,\mathrm{m}$, $\mu = \mu_o$, and $\epsilon = 2\epsilon_o$. Find the guided power flow in watts and the total scattered power of a fiber with a length of $1\,\mathrm{km}$. The ratio of the scattered power to the guided power gives the intrinsic loss of the fiber, a lower bound no matter how pure the fiber. Estimate, with the numbers given above, the fiber loss per kilometer (in dB/km) due to the Rayleigh scattering.

(e) The scattering of microwave by rain drops is another example of Rayleigh scattering. Assume a plane wave at a frequency of 10 GHz is incident on a rain drop of radius 1 mm which has a permittivity $\epsilon + i\sigma/\omega$ where $\sigma = \omega\epsilon$ and $\epsilon = 40\epsilon_o$. In addition to the power scattered by the particle, there is also power absorbed due to the imaginary part $i\sigma/\omega$. The absorbed power can be calculated as

$$P_{diss} = \frac{1}{2}\int dV\, \sigma\, |E_0|^2$$

Which is the primary loss of power in the microwave, scattering or absorption?

Problem P6.3

Consider scattering by a conducting cylinder of radius a of an incident wave with magnetic field $\overline{H}$ parallel to the axis of the cylinder. Find the scattered field in the limit $ka \ll 1$.

Problem P6.4

Prove (6.2).

Problem P6.5

Calculate $\overline{\overline{Q}}_D^{\pm}$, $\overline{\overline{Q}}_N^{\pm}$, $\overline{\overline{Q}}_{D1}^{\pm}$, and $\overline{\overline{Q}}_{N1}^{\pm}$ for the following periodic surface height:

$$f(x) = \begin{cases} h/2 & 0 \le x < p/2 \\ -h/2 & p/2 \le x < p \end{cases}$$

Problem P6.6

Consider a Gaussian random process $f(x,y)$ with $< f >= 0$ and

$$< f(x_1,y_1)\ f(x_2,y_2) >= \sigma^2\, e^{-[(x_1-x_2)^2+(y_1-y_2)^2]/L^2}$$

Letting $\alpha = \sigma f(x,y)/\sigma x$ and $\beta = \sigma f(x,y)/\sigma y$, calculate .

$$< \alpha(x_1,y_1)\alpha(x_2,y_2) >,\ < \beta(x_1,y_1)\beta(x_2,x_2) >$$

and

$$< \alpha(x_1,y_1)\beta(x_2,y_2) >$$

Problem P6.7

Let $f(x)$ be a stationary Gaussian random process with variance σ^2 and $< f(x) >= 0$. Show that

$$< e^{i\alpha f(x)} > \ = e^{-\sigma^2\alpha^2/2}$$

Problem P6.8

Let $f(x)$ be a stationary Gaussian random process with $< f(x) > = 0$ and

$$< f(x_1)f(x_2) >= \sigma^2 e^{-(x_1-x_2)^2/l^2}$$

Calculate $< e^{i\alpha[f(x_1)-f(x_2)]} >$.

Problem P6.9

For a Gaussian random rough surface, show that

$$p(\alpha,\beta) = \frac{e^{-(\alpha^2+\beta^2)/2\sigma^2|c''(0)|}}{2\pi\sigma^2|c''(0)|}$$

Problem P6.10

Consider the bistatic scattering coefficients for a Gaussian random rough surface $f(x,y)$ under the geometrical optics approximation. Express your answer in terms of the slope probability density function $p(\alpha,\beta)$, where $\alpha = \partial f/\partial x$ and $\beta = \partial f/\partial y$.

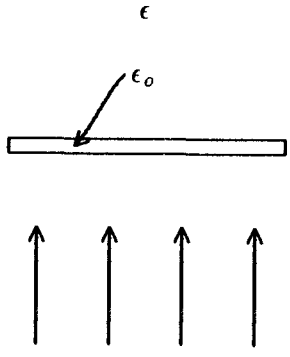


Figure P6.12a External field applied to a disk-shaped cavity.

Problem P6.11

Plot the backscattering coefficients for a random rough surface for $\sigma/L = 0.1$ as a function of incidence angle θ_i using the geometric optics approximation. Let $\epsilon_1 = 80\epsilon_o$.

Problem P6.12

We shall study the behavior of the field inside a cavity in a polarizable medium with an externally applied electrostatic field. Assume that the applied field is along the $\hat{z}$ direction, i.e., $\overline{E}_a = \hat{z}E_0$ and the polarization vector is $\overline{P}$. Show that

(i) If the shape of the cavity is a sphere, the field inside is $\overline{E} = \overline{E}_a + \overline{P}/3\epsilon_o$.

(ii) If the shape of the cavity is a disk [Fig. P6.12a], a limiting case of an ellipsoid perpendicular to the applied field), the field inside is $\overline{E} = \overline{E}_a + \overline{P}/\epsilon_o$.

(iii) If the shape of the cavity is of a needle shape (a limiting case of the two axes of an ellipsoid approaching zero), which is aligned in the same direction as the applied field [Fig. P6.12b], the field inside the cavity is $\overline{E} = \overline{E}_a$.

The field inside a cavity is related to the singular behavior of the dyadic Green's function when the observation point is in the source region. To overcome such a difficulty, a ***principal value*** integral is defined. Thus, the electric field when the observation point is in the source

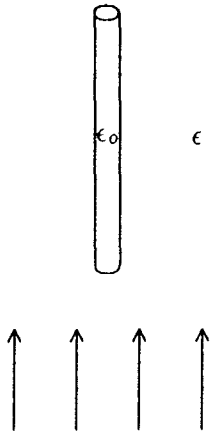


Figure P6.12b External field applied to needle-shaped cavity.

region is given as

$$\overline{E}(r) = i\omega\mu_o \,\mathrm{PV} \int \overline{\overline{G}}(r,r') \cdot \overline{J}(r')\, dV' + \frac{\overline{\overline{L}} \cdot \overline{J}(r)}{i\omega\epsilon_o}$$

where the principal value integral is obtained by performing the volume integration over the source region excluding a small cavity called *exclusion volume* with a certain shape and containing the observation point. The size of the cavity is allowed to shrink to zero eventually. Thus, the first term can be thought of as the contribution to the field at the observation point due to the source outside the exclusion volume. The latter term is the contribution to the field at the observation point due to the source inside the exclusion volume, which can be shown to be non-vanishing event when the exclusion volume shrinks to zero. In this definition, $\overline{E}(r)$ is unique whereas the value of each of the two terms on the right-hand side of the equation depends on the shape of the exclusion volume chosen. This is evident by finding the value of $\overline{\overline{L}}$ for different cavity shapes. Since the exclusion volume is vanishingly small, we can assume that $\overline{J}(r)$ is uniform inside the cavity. Show that

(i) If the exclusion volume is chosen to be a sphere, $\overline{\overline{L}} = \overline{\overline{I}}/3$ where $\overline{\overline{I}}$ is the unit dyad.

(ii) If the exclusion volume is a disk, $\overline{\overline{L}} = \hat{z}\hat{z}$.

(iii) If the exclusion volume is a long cylinder, $\overline{\overline{L}} = \overline{\overline{I}}^t/2$, where $\overline{\overline{I}}^t$ is a unit dyad transverse to the axis of the cylinder. The correspondence of the above with the dielectric medium cavity problem can be noted if we assume $\overline{J}$ to consist of polarization current $\overline{J}_p$ only. In such a case, $\overline{J}_p = -i\omega\overline{P}$. The field inside the cavity is the principal value integral. It is given by $\overline{E}(r) = \overline{\overline{L}} \cdot \overline{P}/\epsilon_o$.

(iv) Explain the apparent disagreement for the long cylinder case.

Problem P6.13

The self-similar model is used to explain the dielectric response of rocks at very low frequency so that attenuation due to scattering can be ignored [Sen et al., 1981].

When a field $\overline{E}^{inc} = \hat{z}$ is incident on a spherical particle with permittivity ϵ_s, the potential outside the particle is

$$\Phi_{out} = -r\cos\theta + \frac{A}{r^2}\cos\theta$$

and

$$A = \frac{\epsilon_s - \epsilon_o}{2\epsilon_o + \epsilon_s} a^3$$

where $v_o = 4\pi a^3/3$ is the volume of the sphere. Thus, when A is calculated, the permittivity is given by

$$\frac{\epsilon_s}{\epsilon_o} = \frac{1 + 2A/a^3}{1 - A/a^3}$$

Consider a sphere of material of permittivity ϵ_m (*matrix*) coated with a material of permittivity ϵ_w (*water*). The inner radius is b and the outer radius is a. An electric field $\overline{E}^{inc} = \hat{z}$ is incident on it. Show that

$$A = a^3 \frac{(\epsilon_w - \epsilon_o)(\epsilon_m + 2\epsilon_w) + \eta(2\epsilon_w + \epsilon_o)(\epsilon_m - \epsilon_w)}{(\epsilon_w + 2\epsilon_o)(\epsilon_m + 2\epsilon_w) + \eta(2\epsilon_w - 2\epsilon_o)(\epsilon_m - \epsilon_w)}$$

Show that the overall permittivity of the coated sphere is

$$\epsilon_{cs} = \epsilon_w \left[\frac{\epsilon_m + 2\epsilon_w + 2\eta(\epsilon_m - \epsilon_w)}{\epsilon_m + 2\epsilon_w - \eta(\epsilon_m - \epsilon_w)} \right]$$

where $\eta = b^3/a^3$ is the volume fraction of the inner sphere.

Using the strong permittivity mixing formula

$$\sum_{p=1}^{n} \frac{\epsilon_p - \epsilon_g}{\epsilon_p + 2\epsilon_g} f_p = 0$$

the effective permittivity of a mixture of n types of coated spheres is

$$\sum_{p=1}^{n} f_p \frac{\epsilon_{cs}^{(p)} - \epsilon_g}{\epsilon_{cs}^{(p)} + 2\epsilon_g} = 0$$

where f_p is the volume fraction of the pth species and $\epsilon_{cs}^{(p)}$ is the corresponding coated sphere permittivity . Let η_p be the volume fraction of the inner sphere of the pth species. Consider the following simple model: all spheres in rocks have the same η. That is, $\eta_p = \eta$ for all $p = 1, 2, \ldots, n$. Hence, $\epsilon_{cs}^{(p)} = \epsilon_{cs}$ for all p. Show that $\epsilon_g = \epsilon_{cs}$. The rock is of porosity ϕ which is occupied by water. Hence $\eta = 1 - \phi$. Thus,

$$\epsilon_g = \epsilon_{cs} = \epsilon_w \left(\frac{\epsilon_m + 2\epsilon_w + 2(1-\phi)(\epsilon_m - \epsilon_w)}{\epsilon_m + 2\epsilon_w - (1-\phi)(\epsilon_m - \epsilon_w)} \right)$$

In the very low frequency limit, we have $\epsilon_m = \epsilon_m' + i\sigma_m/\omega$ and $\epsilon_w = \epsilon_w' + i\sigma_w/\omega$ where σ_m and σ_w are the conductivities, respectively, of rock matrix and water. Let $\epsilon_g = \epsilon_g' + i\sigma_g/\omega$ where σ_g is the effective conductivity. Show that in the usual case when σ_w/ω dominates in the quasistatic limit

$$\sigma_g = \sigma_w \frac{2\phi}{3 - \phi}$$

Thus σ_g is proportional to ϕ for small ϕ. By using a differential approach, the self-similar model further shows that $\sigma_g \approx \phi^{3/2}$ for small ϕ [Sen et al., 1981].

The Maxwell-Garnett mixing formula is

$$\frac{\epsilon - \epsilon_1}{\epsilon + 2\epsilon_1} = f_2 \frac{\epsilon_2 - \epsilon_1}{\epsilon_2 + 2\epsilon_1}$$

where f_2 is the fractional volume of species 2. The fractional volume f_1 is equal to $1 - f_2$. The Maxwell-Garnett formula is derived under the assumption that ϵ_1 is the permittivity of the background medium

and ϵ_2 is that of the scatterer. In the case of rocks, assume that water is the background and the rock grains are the scatterers by letting $\epsilon_1 = \epsilon_w$, $\epsilon_2 = \epsilon_m$ and $f_2 = 1 - \phi$. Solve for ϵ.

Problem P6.14

Assume a mixture of two constituents. One constituent is background medium with permittivity ϵ_b and fractional volume f_b. The other constituent is scatterer with permittivity ϵ_s and fractional volume f_s. The scatterers are assumed to be spherical with radius a. We have $f_b + f_s = 1$. Next use simple physical arguments to construct correlation functions of $\xi(\bar{r})$ for the mixture. First look at the one-dimensional problem. Consider pulses of height ξ_s and width l (which corresponds to $2a$ in the three-dimensional case) distributed on a line. For the rest of the line, the amplitude is ξ_b [Fig. P6.14a]. The problem is to determine the correlation function. To simplify the problem, the line is partitioned into segments of length l, with the amplitude in each partition assuming either the value ξ_s or ξ_b [Fig. P6.14b]. Let $\xi(k)$ be the value of amplitude of partition k. Then, $Pr(\xi(k) = \xi_s) = f_s$, $Pr(\xi(k) = \xi_b) = f_b$ and $< \xi(k) > = f_s\xi_s + f_b\xi_b = 0$. To calculate the correlation function, verify that

$$< \xi(k)\xi(k+m) > = f_s\xi_s^2 + f_b\xi_b^2 \qquad \text{for } m = 0$$

For $m \neq 0$,

$$\begin{aligned} < \xi(k)\xi(k+m) > &= Pr\left[\xi(k) = \xi_s, \xi(k+m) = \xi_s\right] \xi_s^2 \\ &\quad + Pr\left[\xi(k) = \xi_s, \xi(k+m) = \xi_b\right] \xi_s\xi_b \\ &\quad + Pr\left[\xi(k) = \xi_b, \xi(k+m) = \xi_s\right] \xi_b\xi_s \\ &\quad + Pr\left[\xi(k) = \xi_b, \xi(k+m) = \xi_b\right] \xi_b \\ &= f_s^2\xi_s^2 + 2f_bf_s\xi_s\xi_b + f_b^2\xi_b^2 = 0 \end{aligned}$$

Using the results for the three-dimensional continuum case gives

$$R_\xi(r) = \begin{cases} < \xi^2 > & \text{for } 0 \leq r \leq a \\ 0 & \text{otherwise} \end{cases}$$

$$C_\xi(r) = \begin{cases} < |\xi|^2 > & \text{for } 0 \leq r \leq a \\ 0 & \text{otherwise} \end{cases}$$

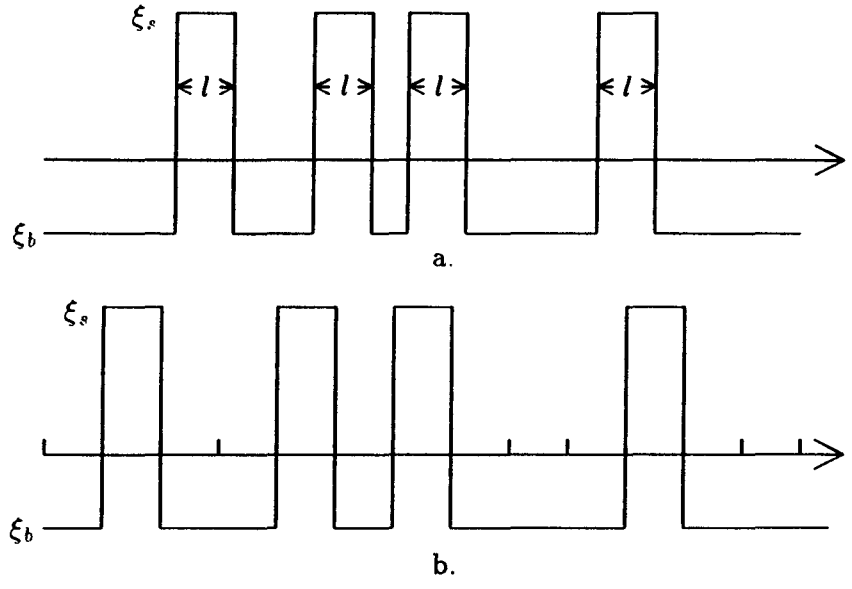


Figure P6.14

Extend the method to treat a mixture of three constituents and derive

$$R_\xi(r) = \begin{cases} <\xi^2> & \text{for } 0 \le r \le a_1 \\ \xi_{s2}^2 f_{s2}/(1-f_{s2}) & \text{for } a_1 < r \le a_2 \\ 0 & \text{otherwise} \end{cases}$$

$$C_\xi(r) = \begin{cases} <|\xi|^2> & \text{for } 0 \le r \le a \\ |\xi_{s2}|^2 f_{s2}/(1-f_{s2}) & \text{for } a_1 < r \le a_2 \\ 0 & \text{otherwise} \end{cases}$$

Problem P6.15

By using the Fourier transform technique, we express the dyadic Green's function for a two-layer medium as the following integral:

$$\overline{\overline{G}}_{01}(\bar{r},\bar{r}') = \iint d\bar{k}_\perp \, \overline{\overline{g}}_{01}(\bar{k}_\perp, z, z') e^{i\bar{k}_\perp \cdot (\bar{r}_\perp - \bar{r}'_\perp)}$$

where

$$\bar{\bar{g}}_{01}(\bar{k}_\perp, z, z') = \frac{i}{8\pi^2}\frac{e^{ik_z z}}{k_{1z}}$$
$$\left[\frac{T_{10}^{\mathrm{TE}}}{D_2(k_\perp)}\hat{e}(k_{0z})\left\{R_{12}^{\mathrm{TE}}\, e^{i2k_{1z}d}\,\hat{e}_1(-k_{1z})e^{ik_{1z}z'} + \hat{e}_1(k_{1z})e^{-ik_{1z}z'}\right\}\right.$$
$$\left. + \frac{k_1}{k}\frac{T_{10}^{\mathrm{TM}}}{F_2(k_\perp)}\hat{h}(k_{0z})\left\{R_{12}^{\mathrm{TM}}\, e^{i2k_{1z}d}\,\hat{h}_1(-k_{1z})e^{ik_{1z}z'} + \hat{h}_1(k_{1z})e^{-ik_{1z}z'}\right\}\right]$$

In the far-field limit, using a two-dimensional stationary-phase method to show that as $kr \to \infty$,

$$e^{i\bar{k}\cdot\bar{r}} = e^{i(k_x x + k_y y + \sqrt{k^2 - k_x^2 - k_y^2}\, z)}$$

With the observation point at $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, and $z = r\cos\theta$, Show that the stationary-phase point is specified by

$$k_x = k\sin\theta\cos\phi$$
$$k_y = k\sin\theta\sin\phi$$

Carry out the necessary steps to obtain

$$\bar{\bar{G}}_{01}(\bar{r}, \bar{r}') = \frac{e^{ikr}}{4\pi r}\left\{\left[\frac{T_{01}^{\mathrm{TE}}}{D_2}\,\hat{e}(k_{0z})\hat{e}_1(k_{1z}) + \frac{k}{k_1}\frac{T_{01}^{\mathrm{TM}}}{F_2}\hat{h}(k_{0z})\hat{h}_1(k_{1z})\right]e^{-i\bar{k}_1\cdot\bar{r}'}\right.$$
$$+ \left[\frac{T_{01}^{\mathrm{TE}}}{D_2}R_{12}^{\mathrm{TE}}\,\hat{e}(k_{0z})\hat{e}_1(-k_{1z}) + \frac{k}{k_1}\frac{T_{01}^{\mathrm{TM}}}{F_2}R_{12}^{\mathrm{TM}}\,\hat{h}(k_{0z})\hat{h}_1(-k_{1z})\right]$$
$$\left. \cdot\, e^{i2k_{1z}d - i\bar{k}_1\cdot\bar{r}'}\right\}$$

with all k_x and k_y equal to the values at the stationary-phase point.

Problem P6.16

Considering matter as composed of spherical molecules immersed in free space, the Clausius-Mossotti formula for the effective permittivity is

$$\epsilon = \epsilon_o\frac{1 + 2N\alpha/3\epsilon_o}{1 - N\alpha/3\epsilon_o}$$

where

$N =$ number of molecules per unit volume
$\alpha =$ constant defined by $\overline{P} = N\alpha\overline{E}_{local}$.

(a) To determine the constant α, we first calculate the polarization $\overline{P}$ of a spherical molecule with dielectric constant ϵ_s in a uniform static field $\overline{E}_{local}$. It is given by

$$\overline{E}_{local} = \overline{E}^{in} + \frac{\overline{P}}{3\epsilon_o}$$

where $\overline{E}^{in}$ is the field inside the sphere. From the expression of $\overline{E}^{in}$, deduce that

$$\overline{P} = 3\epsilon_o \frac{\epsilon_s - \epsilon_o}{\epsilon_s + 2\epsilon_o} \overline{E}_{local}$$

Let f be the fractional volume of the molecules, or

$$f = \frac{\text{total volume of molecules}}{\text{total volume of the material}}$$

$$y = \frac{\epsilon_s - \epsilon_o}{\epsilon_s + 2\epsilon_o}$$

and

$$v_o = \frac{4}{3}\pi a^3$$

prove the Maxwell-Garnett mixing formula:

$$\epsilon_{eff} = \epsilon_o \frac{1 + 2fy}{1 - fy}$$

(b) When the frequency gets higher, the loss due to scattering becomes significant and we need to include an imaginary term in the effective permittivity. For simplicity, apply the Rayleigh scattering model. Show that the effective permittivity is

$$\epsilon_{eff} = \epsilon_o \left\{ \frac{1 + 2fy}{1 - fy} + i2fk^3a^3y^2 \right\}$$

VII

ELECTROMAGNETIC WAVE THEORY AND SPECIAL RELATIVITY

7.1 Lorentz Transformation
7.2 Maxwell-Minkowski Theory
  a. Minkowski Formulation
  b. Amperian Formulation
  c. Chu Formulation
7.3 Derivation of Transformation Formulas
  a. Electromagnetic Field Transformation
  b. Electromagnetic Field Classification
7.4 Transformation of Constitutive Relations
  a. Moving Isotropic Media
  b. Moving Bianisotropic Media
  c. Moving Gyrotropic Media
  d. Moving Uniaxial Media
  e. Accelerated Media
7.5 Transformation of Frequency and Wave Vector
  a. Aberration Effect
  b. Doppler Effect
7.6 Plane Waves in Moving Uniaxial Media
7.7 Phase Matching at Moving Boundaries
7.8 Guided Waves in a Moving Dielectric Slab
7.9 Guided Waves in Moving Gyrotropic Media
7.10 Four-Dimensional Notations
  a. Contravariant and Covariant Vectors
  b. Maxwell's Equations in Tensor Form
  c. Constitutive Relations in Tensor Form
7.11 Hamilton's Principle and Noether's Theorem
  a. Action Integral
  b. Hamilton's Principle and Maxwell's Equations
  c. Noether's Theorem and Energy Momentum Tensors
Problems

7.1 Lorentz Transformation

Electromagnetic theory was the most important stimulus that prompted Einstein's special theory of relativity. The principle of relativity, which requires that all physical laws be form-invariant as described by all observers, is basic in formulating the laws of nature. Space and time constitute the coordinates for descriptions of physical phenomena. The Galilean transformation of space and time was used to provide a basis for deriving transformation laws between observers in relative motion. The principle of relativity that is based on the Galilean transformation is referred to as *Galilean relativity*. Under the Galilean transformation, the laws of Newtonian mechanics are form-invariant, but the laws of electromagnetism change their form. In 1904, Lorentz examined the conditions for form invariance of Maxwell's equations in vacuum between observers moving with constant velocities relative to each other. In 1905, Einstein deduced Lorentz transformation laws from the single postulate that the velocity of light in vacuum is a universal constant, and the assumption that vacuum is linear, isotropic, and homogeneous. Einstein's principle of relativity that is based on the Lorentz transformation is *special relativity*. The laws of Newtonian mechanics, since they were not form-invariant under the Lorentz transformation, have been revised. Physical laws that are form-invariant under the Lorentz transformation are *Lorentz-covariant*.

Consider the simple case in which the coordinate axes of observers S and S' are parallel, with their origins coinciding at time $t = 0$. Observer S' moves uniformly with velocity $\overline{v}$ relative to S [Fig. 7.1.1]. The Lorentz transformation of space-time coordinates between these two moving observers, with the use of dyadic notation, is given by

$$\mathrm{LT}\begin{cases} ct' = \gamma ct - \gamma\overline{\beta}\cdot\overline{r} & (1a) \\ \overline{r}' = \overline{\overline{\alpha}}\cdot\overline{r} - \gamma\overline{\beta}ct & (1b) \end{cases}$$

where

$$\overline{\overline{\alpha}} = \overline{\overline{I}} + (\gamma - 1)\frac{\overline{\beta}\,\overline{\beta}}{\beta^2} \tag{2}$$

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{3}$$

$$\beta^2 = \overline{\beta}\cdot\overline{\beta} \qquad \overline{\beta} = \frac{\overline{v}}{c} \tag{4}$$

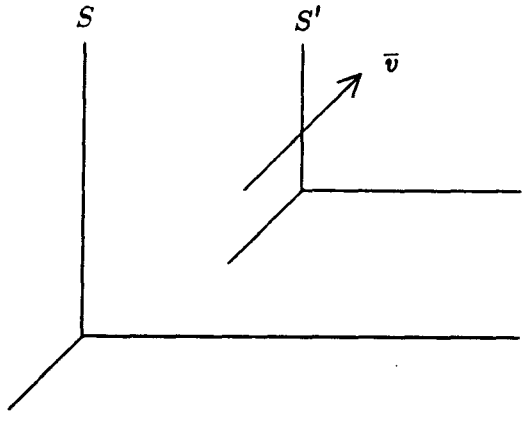


Figure 7.1.1 Observer S' moves with velocity $\bar{v}$ relative to observer S.

and $c = 3 \times 10^8$ m/s is the velocity of light in vacuum. In matrix notation, the unit dyad $\bar{\bar{I}}$ is a diagonal matrix

$$\bar{\bar{I}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and

$$\bar{\bar{\alpha}} = \begin{bmatrix} 1+(\gamma-1)\dfrac{\beta_x\beta_x}{\beta^2} & (\gamma-1)\dfrac{\beta_x\beta_y}{\beta^2} & (\gamma-1)\dfrac{\beta_x\beta_z}{\beta^2} \\ (\gamma-1)\dfrac{\beta_y\beta_x}{\beta^2} & 1+(\gamma-1)\dfrac{\beta_y\beta_y}{\beta^2} & (\gamma-1)\dfrac{\beta_y\beta_z}{\beta^2} \\ (\gamma-1)\dfrac{\beta_z\beta_x}{\beta^2} & (\gamma-1)\dfrac{\beta_z\beta_y}{\beta^2} & 1+(\gamma-1)\dfrac{\beta_z\beta_z}{\beta^2} \end{bmatrix}$$

where β_x, β_y, and β_z are the x, y, and z components of $\bar{\beta}$. Clearly, because $\bar{\bar{\alpha}}$ is symmetrical, the position vector $\bar{r} = \hat{x}x + \hat{y}y + \hat{z}z$ can be viewed as a column matrix operated on by $\bar{\bar{\alpha}}$, giving rise to another column matrix.

When $\bar{v}$ is along the z axis, $\bar{\beta} = \hat{z}\beta$, $\beta_x = \beta_y = 0$, and $\beta_z = \beta$. The $\bar{\bar{\alpha}}$ matrix simplifies to

$$\bar{\bar{\alpha}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \gamma \end{bmatrix}$$

We find from (1)

$$ct' = \gamma(ct - \beta z) \tag{5a}$$

$$x' = x \tag{5b}$$

$$y' = y \tag{5c}$$

$$z' = \gamma(z - \beta ct) \tag{5d}$$

We observe that the time coordinate is not a universal constant; two physical events that are simultaneous in S' will no longer be simultaneous in S.

An important identity can be derived from the LT (1). Forming the difference of magnitude squares of $\bar{r}'$ and ct', and using (1), we find

$$|\bar{r}'|^2 - |ct'|^2 = |\bar{r}|^2 - |ct|^2 \tag{6}$$

Equation (6) is important because it is independent of the relative velocity $\bar{v}$ between S and S'. It is a numerical constant that is invariant under the Lorentz transformation. Its square root can be regarded as expressing the length of a four-dimensional vector representing the space and time coordinates of a physical event. Evidently, in this four-dimensional space, called *Minkowski space*, the length of a vector can be imaginary as well as real.

When $\bar{v}$ is so small that only terms of the order of $\bar{v}/c$ are significant, we have $\bar{\bar{\alpha}} = \bar{\bar{I}}$, $\gamma = 1$, and LT (1) becomes

$$\text{FOLT} \begin{cases} ct' = ct - \bar{\beta} \cdot \bar{r} & (7a) \\ \bar{r}' = \bar{r} - \bar{\beta} ct & (7b) \end{cases}$$

We call (7) the *First-Order Lorentz Transformation* (FOLT). As is seen from FOLT, the space term $\beta \cdot \bar{r}$ in the time transformation may not be negligible, since $\bar{r}$ and $\bar{r}'$ can be large, while $\bar{\beta}$ is small.

Before 1905, time was regarded as a universal quantity. For two observers in relative uniform motion, the space coordinate changed because of motion, but the time coordinate remained the same:

$$\text{GT} \begin{cases} t' = t \\ \bar{r}' = \bar{r} - \bar{v}t \end{cases}$$

This transformation law of space and time is the Galilean transformation (GT). We note also that the Galilean transformation is not a

limiting case of the LT when velocity $\overline{v}$ is small. Rather, LT reduces to GT when $\overline{v}$ is small and when $\overline{r}$ is small compared with ct/β, as is seen from FOLT $(7a)$. Mathematically, LT reduces to GT when $c \to \infty$.

The transformation laws for space and time under GT, LT, and FOLT have been given from S to S'. The transformations from S' to S can be obtained by finding their inverses. Taking, for example, the LT, we have

$$\text{LT} \begin{cases} ct = \gamma ct' + \gamma \overline{\beta} \cdot \overline{r}' & (8a) \\ \overline{r} = \overline{\overline{\alpha}} \cdot \overline{r}' + \gamma \overline{\beta} ct' & (8b) \end{cases}$$

The FOLT follows from (8) by letting $\gamma = 1$, and the GT follows from (8) by letting $c \to \infty$. Physically, the formulas can be obtained by changing $\overline{\beta}$ in the transformation laws to $-\overline{\beta}$; as S' is moving with $\overline{v}$ relative to S, S is moving with $-\overline{v}$ relative to S'.

7.2 Maxwell-Minkowski Theory

For the principle of special relativity, we can state the two postulates as follows: (i) among uniformly moving observers, space and time coordinates obey Lorentz transformation laws; and (ii) physical laws are form-invariant under the Lorentz transformation laws for space and time. On the basis of these two postulates, the constancy of the velocity of light is a direct consequence of the Lorentz covariance of Maxwell's equations in vacuum. In 1908 Minkowski postulated that the macroscopic Maxwell's equations in material are Lorentz covariant. With this postulate and the Lorentz transformation for space and time cordinates, we can obtain the transformation formulas for electromagnetic field vectors, from which the constitutive relations for various moving media can be derived.

a. Minkowski Formulation

Suppose that, from the point of view of an observer S, macroscopic electrodynamics is described by Maxwell's equations:

$$\nabla \times \overline{E} + \frac{\partial \overline{B}}{\partial t} = 0 \qquad (1)$$

$$\nabla \cdot \overline{B} = 0 \qquad (2)$$

$$\nabla \times \overline{H} - \frac{\partial \overline{D}}{\partial t} = \overline{J} \tag{3}$$

$$\nabla \cdot \overline{D} = \rho \tag{4}$$

with the current-charge conservation law

$$\nabla \cdot \overline{J} + \frac{\partial \rho}{\partial t} = 0 \tag{5}$$

Then, from the point of view of an observer S' moving with respect to S, Maxwell's equations take the same form:

$$\nabla' \times \overline{E}' + \frac{\partial \overline{B}'}{\partial t'} = 0 \tag{6}$$

$$\nabla' \cdot \overline{B}' = 0 \tag{7}$$

$$\nabla' \times \overline{H}' - \frac{\partial \overline{D}'}{\partial t'} = \overline{J}' \tag{8}$$

$$\nabla' \cdot \overline{D}' = \rho' \tag{9}$$

with the current-charge conservation law

$$\nabla' \cdot \overline{J}' + \frac{\partial \rho'}{\partial t'} = 0 \tag{10}$$

where primes denote quantities associated with S'. The fundamental field quantities are $\overline{E}$, $\overline{B}$, $\overline{D}$, and $\overline{H}$. If $\overline{E}$ and $\overline{B}$ are regarded as pure field quantities, then $\overline{D}$ and $\overline{H}$ contain information about the material media. Following Sommerfeld, we refer to $\overline{E}$ and $\overline{B}$ as the entities of intensity, and $\overline{D}$ and $\overline{H}$ as the entities of quantity. In four-dimensional Minkowski space the entities of intensity form a field tensor of second rank, and the entities of quantity form an excitation tensor of second rank. By Minkowski's postulate, we can find the transformation laws for all field variables from the Lorentz transformation of space and time.

The formulation that we just described is called the *Minkowski formulation*. The concept that all material media can be regarded as source terms in Maxwell's equations and that only two electromagnetic field vectors are fundamental quantities has led to alternative formulations for macroscopic electromagnetic theory.

b. Amperian Formulation

In the ***Amperian formulation***, Maxwell's equations take the form

$$\nabla \times \overline{E}_A = -\frac{\partial \overline{B}_A}{\partial t}$$

$$\nabla \times \left[\frac{1}{\mu_o}\overline{B}_A\right] = \frac{\partial}{\partial t}\left(\epsilon_o \overline{E}_A + \overline{P}_A - \frac{1}{c^2}\overline{M}_A \times \overline{v}\right) + \nabla \times \left(\overline{M}_A + \overline{P}_A \times \overline{v}\right) + \overline{J}_A$$

$$\nabla \cdot (\epsilon_o \overline{E}_A) = -\nabla \cdot \left(\overline{P}_A - \frac{1}{c^2}\overline{M}_A \times \overline{v}\right) + \rho_A$$

$$\nabla \cdot \overline{B}_A = 0$$

where the subscript A signifies the Amperian formulation. The two fundamental field vectors are $\overline{E}_A$ and $\overline{B}_A$, while the polarization vector $\overline{P}_A$ and the magnetization $\overline{M}_A$ characterize the material media moving with velocity $\overline{v}$. The variables are related to those in the Minkowski formulation by

$$\overline{E} = \overline{E}_A$$

$$\overline{B} = \overline{B}_A$$

$$\overline{H} = \frac{1}{\mu_o}\overline{B}_A - \overline{M}_A - \overline{P}_A \times \overline{v}$$

$$\overline{D} = \epsilon_o \overline{E}_A + \overline{P}_A - \frac{1}{c^2}\overline{M}_A \times \overline{v}$$

$$\overline{J} = \overline{J}_A$$

$$\rho = \rho_A$$

In the ***Boffi formulation***, Maxwell's equations take the form

$$\nabla \times \overline{E}_B = -\frac{\partial \overline{B}_B}{\partial t}$$

$$\nabla \times \left(\frac{1}{\mu_o}\overline{B}_B\right) = \frac{\partial}{\partial t}\left(\epsilon_o \overline{E}_B + \overline{P}_B\right) + \nabla \times \overline{M}_B + \overline{J}_B$$

$$\nabla \cdot \epsilon_o \overline{E}_B = -\nabla \cdot \overline{P}_B + \rho_B$$

$$\nabla \cdot \overline{B}_B = 0$$

where the subscript B signifies the Boffi formulation. The variables are related to those in the Minkowski formulation by

$$\begin{aligned}
\overline{E} &= \overline{E}_B \\
\overline{B} &= \overline{B}_B \\
\overline{H} &= \frac{1}{\mu_o}\overline{B}_B - \overline{M}_B \\
\overline{D} &= \epsilon_o \overline{E}_B + \overline{P}_B \\
\overline{J} &= \overline{J}_B \\
\rho &= \rho_B
\end{aligned}$$

c. *Chu Formulation*

In the ***Chu formulation***, Maxwell's equations take the form

$$\begin{aligned}
\nabla \times \overline{E}_C &= -\frac{\partial}{\partial t}\left(\mu_o \overline{H}_C + \mu_o \overline{M}_C\right) - \nabla \times \left(\mu_o \overline{M}_C \times \overline{v}\right) \\
\nabla \times \overline{H}_C &= \frac{\partial}{\partial t}\left(\epsilon_o \overline{E}_C + \overline{P}_C\right) + \nabla \times \left(\overline{P}_C \times \overline{v}\right) + \overline{J}_C \\
\nabla \cdot \epsilon_o \overline{E}_C &= -\nabla \cdot \overline{P}_C + \rho_C \\
\nabla \cdot \mu_o \overline{H}_C &= -\nabla \cdot \mu_o \overline{M}_C
\end{aligned}$$

where the subscript C signifies the Chu formulation. The variables are related to those in the Minkowski formulation by

$$\begin{aligned}
\overline{E} &= \overline{E}_C + \mu_o \overline{M}_C \times \overline{v} \\
\overline{B} &= \mu_o \overline{H}_C + \mu_o \overline{M}_C \\
\overline{H} &= \overline{H}_C - \overline{P}_C \times \overline{v} \\
\overline{D} &= \epsilon_o \overline{E}_C + \overline{P}_C \\
\overline{J} &= \overline{J}_C \\
\rho &= \rho_C
\end{aligned}$$

Maxwell's equations as presented in the various formulations are in indefinite form; constitutive relations for material media have to be supplied. Once the constitutive relations are given, it has been shown by Tai [1964] that all formulations are equivalent.

In the Amperian and the Chu formulations, models of the constituents of media are elaborated with kinematic approaches. In the Amperian formulation [Panofsky and Phillips, 1962], constituents of a dipolar medium are visualized as two basic elements, an electric dipole and a current loop. The Amperian model is closely related to the atomic structure, where spinning and orbiting electrons act as current loops. In the Chu formulation [Fano, Chu, and Adler, 1960], a dipolar medium is visualized as containing electric and magnetic dipoles. The Chu formulation is useful because there are no inherently moving parts in a magnetic dipole, as opposed to a current loop. When moments of higher order than the dipole moment are significant in a medium, the task of modeling becomes much more involved. From the point of view of electromagnetic wave theory, we are not interested in the reaction of a medium under the action of a field where a model of the medium constituents may be helpful, but we are interested in the way the electromagnetic wave behaves. We favor the Maxwell-Minkowski theory not only because of its simplicity and elegance but also because of its practical applicability.

7.3 Derivation of Transformation Formulas

Transformation formulas for field vectors are direct consequences of the Lorentz transformation for space and time and Minkowski's postulate of the Lorentz covariance of Maxwell's equations. From the LT given in (1) we obtain Lorentz transformation for space–time derivatives. We make use of the chain rule in differentiation,

$$\frac{\partial}{\partial ct} = \left[\frac{\partial ct'}{\partial ct}\right] \frac{\partial}{\partial ct'} + \left[\frac{\partial x_i'}{\partial ct}\right] \frac{\partial}{\partial x_i'}$$

and

$$\frac{\partial}{\partial x_i} = \left[\frac{\partial ct'}{\partial x_i}\right] \frac{\partial}{\partial ct'} + \left[\frac{\partial x_j'}{\partial x_i}\right] \frac{\partial}{\partial x_j'}$$

Substituting the LT (7.1.1) and noting the fact that $\overline{\overline{\alpha}}$ is symmetrical, we find

$$\frac{\partial}{\partial ct} = \gamma \frac{\partial}{\partial ct'} - \gamma \overline{\beta} \cdot \nabla' \tag{1a}$$

$$\nabla = \overline{\overline{\alpha}} \cdot \nabla' - \gamma \overline{\beta} \frac{\partial}{\partial ct'} \tag{1b}$$

To derive transformation laws for all field vectors, we substitute (1) into Maxwell's equations in the S frame and require them to have the same forms in the S' frame. First, consider the charge conservation equation. Transformation from S to S' gives

$$\left[\overline{\overline{\alpha}}\cdot\nabla' - \gamma\overline{\beta}\frac{\partial}{\partial ct'}\right]\cdot\overline{J} + \gamma\left[\frac{\partial}{\partial ct'} - \overline{\beta}\cdot\nabla'\right]c\rho = 0$$

Thus the charge conservation equation is Lorentz-covariant if

$$c\rho' = \gamma(c\rho - \overline{\beta}\cdot\overline{J}) \tag{2a}$$

$$\overline{J}' = \overline{\overline{\alpha}}\cdot\overline{J} - \gamma\overline{\beta}c\rho \tag{2b}$$

Note that a charge distribution stationary in S certainly produces a current in S', but from (2b) a uniform current element in S also generates a charge distribution in S', which is a relativistic effect and cannot be seen under GT.

Next, we introduce (1) into Ampere's law and Gauss' electric field law:

$$\left[\overline{\overline{\alpha}}\cdot\nabla' - \gamma\frac{\partial}{\partial ct'}\overline{\beta}\right]\times\overline{H} - \left[\gamma\frac{\partial}{\partial ct'} - \gamma\overline{\beta}\cdot\nabla'\right]c\overline{D} = \overline{J} \tag{3a}$$

$$\left[\overline{\overline{\alpha}}\cdot\nabla' - \gamma\frac{\partial}{\partial ct'}\overline{\beta}\right]\cdot c\overline{D} = c\rho \tag{3b}$$

To find transformation laws for $\overline{D}'$ and $\overline{H}'$, we wish to cast (3) into the form

$$\nabla'\cdot\overline{D}' = \rho' \tag{4a}$$

$$\nabla'\times\overline{H}' - \frac{\partial}{\partial t'}\overline{D}' = \overline{J}' \tag{4b}$$

In view of (2a), we have $\gamma(3b) - \gamma\overline{\beta}\cdot(3a) = c\rho'$, which gives

$$\gamma\left[(\overline{\overline{\alpha}}\cdot\nabla')\cdot c\overline{D} - \overline{\beta}\cdot(\overline{\overline{\alpha}}\cdot\nabla')\times\overline{H} - \gamma(\overline{\beta}\cdot\nabla')(\overline{\beta}\cdot cD)\right] = c\rho'$$

Using

$$\begin{aligned}\overline{\beta}\cdot[(\overline{\overline{\alpha}}\cdot\nabla')\times\overline{H}] &= \overline{\beta}\cdot\left\{\left[\nabla' + (\gamma-1)\frac{(\overline{\beta}\cdot\nabla')}{\beta^2}\overline{\beta}\right]\times\overline{H}\right\}\\ &= \overline{\beta}\cdot\nabla'\times\overline{H} = -\nabla'\cdot(\overline{\beta}\times\overline{H})\end{aligned}$$

we find

$$\nabla' \cdot \left\{ \gamma \left[\overline{\overline{\alpha}} - \gamma \overline{\beta}\, \overline{\beta} \right] \cdot c\overline{D} + \gamma \overline{\beta} \times \overline{H} \right\} = c\rho' \tag{5a}$$

By the same token, we use (2b) to calculate $\overline{\overline{\alpha}} \cdot (3a) - \gamma \overline{\beta} \cdot (3b) = \overline{J}'$ which gives

$$\begin{aligned} &\overline{\overline{\alpha}} \cdot (\overline{\overline{\alpha}} \cdot \nabla') \times \overline{H} + \gamma \overline{\overline{\alpha}} \cdot (\overline{\beta} \cdot \nabla') c\overline{D} - \gamma \overline{\beta} (\overline{\overline{\alpha}} \cdot \nabla') c\overline{D} \\ &\qquad + \frac{\partial}{\partial ct'} \left[-\gamma \overline{\overline{\alpha}} \cdot (\overline{\beta} \times \overline{H}) - \gamma (\overline{\overline{\alpha}} \cdot c\overline{D}) + \gamma^2 \overline{\beta}\, \overline{\beta} \cdot c\overline{D} \right] = \overline{J}' \end{aligned}$$

Using the fact that

$$\begin{aligned} \overline{\overline{\alpha}} \cdot \left[(\overline{\overline{\alpha}} \cdot \nabla') \times \overline{H} \right] &= \nabla \times \overline{H} + \frac{\gamma - 1}{\beta^2} \\ &\qquad \cdot \left[(\overline{\beta} \cdot \nabla') \overline{\beta} \times \overline{H} - \overline{\beta} \nabla' \cdot (\overline{\beta} \times \overline{H}) \right] \\ &= \nabla' \times \left\{ \overline{H} - \frac{\gamma - 1}{\beta^2} \left[\overline{\beta} \times (\overline{\beta} \times \overline{H}) \right] \right\} \\ &= \nabla' \times \left\{ \gamma \left[\overline{\overline{I}} + \left(\frac{1}{\gamma} - 1 \right) \frac{\overline{\beta}\, \overline{\beta}}{\beta^2} \right] \cdot \overline{H} \right\} \end{aligned}$$

and

$$\begin{aligned} &\overline{\overline{\alpha}} \cdot (\overline{\beta} \cdot \nabla') c\overline{D} - \overline{\beta} \left[(\overline{\overline{\alpha}} \cdot \nabla') c\overline{D} \right] \\ &\qquad = (\overline{\beta} \times \nabla') c\overline{D} - \overline{\beta} (\nabla' \cdot c\overline{D}) \\ &\qquad = -\nabla' \times (\overline{\beta} \times c\overline{D}) \end{aligned}$$

We find

$$\begin{aligned} &\nabla' \times \left\{ \gamma + \left[\overline{\overline{I}} + \left(\frac{1}{\gamma} - 1 \right) \frac{\overline{\beta}\, \overline{\beta}}{\beta^2} \right] \cdot \overline{H} - \gamma \overline{\beta} \times c\overline{D} \right\} \\ &\qquad - \frac{\partial}{\partial ct'} \left\{ \gamma \left[\overline{\overline{\alpha}} - \gamma \overline{\beta}\, \overline{\beta} \right] \cdot c\overline{D} + \gamma \overline{\beta} \times \overline{H} \right\} = \overline{J}' \end{aligned}$$

Comparing (4) and (5), we obtain the transformation formulas for $\overline{D}$ and $\overline{H}$:

$$c\overline{D}' = \gamma \left[\overline{\overline{\alpha}} - \gamma \overline{\beta}\, \overline{\beta} \right] \cdot c\overline{D} + \gamma \overline{\beta} \times \overline{H} \tag{6a}$$

$$\overline{H}' = \gamma \left[\overline{\overline{I}} + \left(\frac{1}{\gamma} - 1 \right) \frac{\overline{\beta}\, \overline{\beta}}{\beta^2} \right] \cdot \overline{H} - \gamma \overline{\beta} \times c\overline{D} \tag{6b}$$

It is easily shown that the dyadic quantities in the square brackets of (6*a*) and (6*b*) are equal to the inverse of $\overline{\overline{\alpha}}$

$$\overline{\overline{\alpha}}^{-1} = \overline{\overline{I}} + \left(\frac{1}{\gamma} - 1\right)\frac{\overline{\beta}\,\overline{\beta}}{\beta^2} = \overline{\overline{\alpha}} - \gamma\overline{\beta}\,\overline{\beta} \tag{7}$$

This is verified by showing that

$$\overline{\overline{\alpha}} \cdot \overline{\overline{\alpha}}^{-1} = \left[\overline{\overline{I}} + (\gamma - 1)\frac{\overline{\beta}\,\overline{\beta}}{\beta^2}\right] \cdot \left[\overline{\overline{I}} + \left(\frac{1}{\gamma} - 1\right)\frac{\overline{\beta}\,\overline{\beta}}{\beta^2}\right] = \overline{\overline{I}}$$

also $\overline{\overline{\alpha}}^{-1} \cdot \overline{\overline{\alpha}} = \overline{\overline{I}}$ and

$$\overline{\overline{\alpha}}^{-1} = \overline{\overline{I}} + \left(\frac{1}{\gamma} - 1\right)\frac{\overline{\beta}\,\overline{\beta}}{\beta^2} = \overline{\overline{\alpha}} - \gamma\overline{\beta}\,\overline{\beta}$$

in view of the fact that $1/\gamma^2 = 1 - \beta^2$. We can further define a 3×3 matrix $\overline{\overline{\beta}}$ such that for any vector $\overline{A}$

$$\overline{\overline{\beta}} \cdot \overline{A} \equiv \overline{\beta} \times \overline{A} \tag{8}$$

In explicit matrix form,

$$\overline{\overline{\beta}} = \begin{bmatrix} 0 & -\beta_z & \beta_y \\ \beta_z & 0 & -\beta_x \\ -\beta_y & \beta_x & 0 \end{bmatrix} \tag{9}$$

From $(\overline{\overline{\beta}})^2 \cdot \overline{A} = \overline{\beta} \times (\overline{\beta} \times \overline{A}) = \overline{\beta}\,\overline{\beta} \cdot \overline{A} - \beta^2\overline{A}$ it follows that

$$\overline{\overline{\beta}}^2 = \overline{\beta}\,\overline{\beta} - \beta^2\overline{\overline{I}} \tag{10}$$

Although both $\overline{\overline{\alpha}}$ and $\overline{\overline{\alpha}}^{-1}$ are symmetric, $\overline{\overline{\beta}}$ is skew-symmetric.

a. Electromagnetic Field Transformation

We now write the Lorentz transformation formulas for $c\overline{D}$ and $\overline{H}$ in the following form:

$$\text{LT} \qquad \begin{bmatrix} c\overline{D}' \\ \overline{H}' \end{bmatrix} = \gamma \begin{bmatrix} \overline{\overline{\alpha}}^{-1} & \overline{\overline{\beta}} \\ -\overline{\overline{\beta}} & \overline{\overline{\alpha}}^{-1} \end{bmatrix} \cdot \begin{bmatrix} c\overline{D} \\ \overline{H} \end{bmatrix} \tag{11}$$

In a similar way, we can show that the forms of Faraday's emf law and Gauss' magnetic field law are preserved under LT, provided that $\overline{E}$ and $\overline{B}$ transform as

$$\text{LT} \qquad \begin{bmatrix} \overline{E}' \\ c\overline{B}' \end{bmatrix} = \gamma \begin{bmatrix} \overline{\overline{\alpha}}^{-1} & \overline{\overline{\beta}} \\ -\overline{\overline{\beta}} & \overline{\overline{\alpha}}^{-1} \end{bmatrix} \cdot \begin{bmatrix} \overline{E} \\ c\overline{B} \end{bmatrix} \tag{12}$$

which is identical in form to (11). Transformation formulas (11) and (12) express the fact that $\overline{H}$ and $\overline{D}$ fields transform as an entity (entity of quantity), and so do $\overline{E}$ and $\overline{B}$ (entity of intensity).

It is interesting to see that field components parallel to the velocity are left unchanged:

$$\overline{E}'_{\|} = \overline{E}_{\|} \tag{13a}$$

$$\overline{B}'_{\|} = \overline{B}_{\|} \tag{13b}$$

$$\overline{D}'_{\|} = \overline{D}_{\|} \tag{13c}$$

$$\overline{H}'_{\|} = \overline{H}_{\|} \tag{13d}$$

and the perpendicular components transform as

$$\overline{E}'_{\perp} = \gamma(\overline{E}_{\perp} + \overline{\beta} \times \overline{c}B_{\perp}) \tag{14a}$$

$$c\overline{B}'_{\perp} = \gamma(c\overline{B}_{\perp} - \overline{\beta} \times \overline{E}_{\perp}) \tag{14b}$$

$$c\overline{D}'_{\perp} = \gamma(c\overline{D}_{\perp} + \overline{\beta} \times \overline{H}_{\perp}) \tag{14c}$$

$$\overline{H}'_{\perp} = \gamma(\overline{H}_{\perp} - \overline{\beta} \times c\overline{D}_{\perp}) \tag{14d}$$

This is in contrast to the transformation of space coordinates, where the perpendicular components are left unchanged.

The FOLT and GT of the field quantities can be easily deduced from the LT. In the case of FOLT, let $\gamma = 1$, and we have

$$\text{FOLT} \qquad \begin{bmatrix} \overline{E}' \\ c\overline{B}' \end{bmatrix} = \begin{bmatrix} \overline{\overline{I}} & \overline{\overline{\beta}} \\ -\overline{\overline{\beta}} & \overline{\overline{I}} \end{bmatrix} \cdot \begin{bmatrix} \overline{E} \\ c\overline{B} \end{bmatrix} \tag{15}$$

$$\text{FOLT} \qquad \begin{bmatrix} c\overline{D}' \\ \overline{H}' \end{bmatrix} = \begin{bmatrix} \overline{\overline{I}} & \overline{\overline{\beta}} \\ -\overline{\overline{\beta}} & \overline{\overline{I}} \end{bmatrix} \cdot \begin{bmatrix} c\overline{D} \\ \overline{H} \end{bmatrix} \tag{16}$$

For GT, let $c \to \infty$;

$$\text{GT} \begin{cases} \overline{E}' = \overline{E} + \overline{v} \times \overline{B} & (17a) \\ \overline{B}' = \overline{B} & (17b) \\ \overline{D}' = \overline{D} & (17c) \\ \overline{H}' = \overline{H} - \overline{v} \times \overline{D} & (17d) \end{cases}$$

The inverses of all of these transformation formulas can be written with $\overline{\beta}$ replaced by $-\overline{\beta}$.

According to both LT (14a) and GT (17a), a pure $\overline{B}$ field in S produces an electric field $\overline{E}'$ in S'. Thus a voltage is induced in a moving conductor when its velocity has a component perpendicular to the $\overline{B}$-field lines. According to LT (14b), for a pure $\overline{E}$ field in S, a magnetic field is witnessed from a moving frame. Thus a stationary electron, when viewed from a moving frame, exhibits a magnetic field. But according to GT (17b), $\overline{B}'$ is equal to $\overline{B}$ and a stationary electron in S exhibits no magnetic field in S'.

We denote the 6×6 matrix in (11) and (12) by $\overline{\overline{L}}_6$:

$$\overline{\overline{L}}_6(\overline{\beta}) = \gamma \begin{bmatrix} \overline{\overline{\alpha}}^{-1} & \overline{\overline{\beta}} \\ -\overline{\overline{\beta}} & \overline{\overline{\alpha}}^{-1} \end{bmatrix} \tag{18}$$

When the velocity is along the z axis, the LT matrix $\overline{\overline{L}}_6$ becomes

$$\overline{\overline{L}}_6 = \gamma \begin{bmatrix} 1 & 0 & 0 & 0 & -\beta & 0 \\ 0 & 1 & 0 & \beta & 0 & 0 \\ 0 & 0 & 1/\gamma & 0 & 0 & 0 \\ 0 & \beta & 0 & 1 & 0 & 0 \\ -\beta & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/\gamma \end{bmatrix}. \tag{19}$$

The inverse transformation is determined by the inverse of $\overline{\overline{L}}_6(\overline{\beta})$. It can be verified that

$$\overline{\overline{L}}_6^{-1}(\overline{\beta}) = \overline{\overline{L}}_6(-\beta) = \gamma \begin{bmatrix} \overline{\overline{\alpha}}^{-1} & -\overline{\overline{\beta}} \\ \overline{\overline{\beta}} & \overline{\overline{\alpha}}^{-1} \end{bmatrix} \tag{20}$$

By physical reasoning, the inverse of a pure Lorentz transformation is equivalent to changing the direction of velocity.

Let us explore further some properties of the $\overline{\overline{L}}_6$ matrix. Since $\overline{\overline{\alpha}}$ is symmetric and $\overline{\overline{\beta}}$ skew-symmetric, we have

$$\overline{\overline{L}}_6^T = \gamma \begin{bmatrix} (\overline{\overline{\alpha}}^{-1})^T & (-\overline{\overline{\beta}})^T \\ \overline{\overline{\beta}}^T & (\overline{\overline{\alpha}}^{-1})^T \end{bmatrix} = \overline{\overline{L}}_6 \tag{21}$$

where the superscript T denotes the transpose of the matrix. Thus $\overline{\overline{L}}_6$ is a symmetric 6×6 matrix. We can also show that

$$\overline{\overline{L}}_6^T \cdot \begin{bmatrix} \overline{\overline{I}} & \overline{\overline{0}} \\ \overline{\overline{0}} & -\overline{\overline{I}} \end{bmatrix} \cdot \overline{\overline{L}}_6 = \begin{bmatrix} \overline{\overline{I}} & \overline{\overline{0}} \\ \overline{\overline{0}} & -\overline{\overline{I}} \end{bmatrix} \tag{22}$$

$$\overline{\overline{L}}_6^T \cdot \begin{bmatrix} \overline{\overline{0}} & \overline{\overline{I}} \\ \overline{\overline{I}} & \overline{\overline{0}} \end{bmatrix} \cdot \overline{\overline{L}}_6 = \begin{bmatrix} \overline{\overline{0}} & \overline{\overline{I}} \\ \overline{\overline{I}} & \overline{\overline{0}} \end{bmatrix} \tag{23}$$

Numerous other identities can also be derived for the $\overline{\overline{L}}_6$ and $\overline{\overline{\alpha}}$ matrices.

Equation (22) can be used to find relations that are invariant under the Lorentz transformation. Using LT for the entity of intensity (12), we have

$$\left[\overline{E}', c\overline{B}'\right] \cdot \begin{bmatrix} \overline{\overline{I}} & \overline{\overline{0}} \\ \overline{\overline{0}} & -\overline{\overline{I}} \end{bmatrix} \cdot \begin{bmatrix} \overline{E}' \\ c\overline{B}' \end{bmatrix} = \left[\overline{E}, c\overline{B}\right] \cdot \overline{\overline{L}}_6^T \cdot \begin{bmatrix} \overline{\overline{I}} & \overline{\overline{0}} \\ \overline{\overline{0}} & -\overline{\overline{I}} \end{bmatrix} \cdot \overline{\overline{L}}_6 \cdot \begin{bmatrix} \overline{E} \\ c\overline{B} \end{bmatrix} \tag{24}$$

In view of (21), (24) gives

$$\left|\overline{E}'\right|^2 - \left|c\overline{B}'\right|^2 = \left|\overline{E}\right|^2 - \left|c\overline{B}\right|^2 \tag{25}$$

It can be seen that the relative velocity between observers S and S' does not appear in (25). The difference between the magnitude squared of $\overline{E}$ and the magnitude squared of $c\overline{B}$ is therefore a numerical constant independent of motion. Any quantity that is invariant under LT is called a Lorentz invariant. Note the difference between Lorentz invariance and Lorentz covariance: the former refers to a scalar number, the latter to a physical law. Another Lorentz invariant is obtained by using (12) and (23)

$$\left[\overline{E}', c\overline{B}'\right] \cdot \begin{bmatrix} \overline{\overline{0}} & \overline{\overline{I}} \\ \overline{\overline{I}} & -\overline{\overline{0}} \end{bmatrix} \cdot \begin{bmatrix} \overline{E}' \\ c\overline{B}' \end{bmatrix} = \left[\overline{E}, c\overline{B}\right] \cdot \overline{\overline{L}}_6^T \cdot \begin{bmatrix} \overline{E} \\ c\overline{B} \end{bmatrix}$$

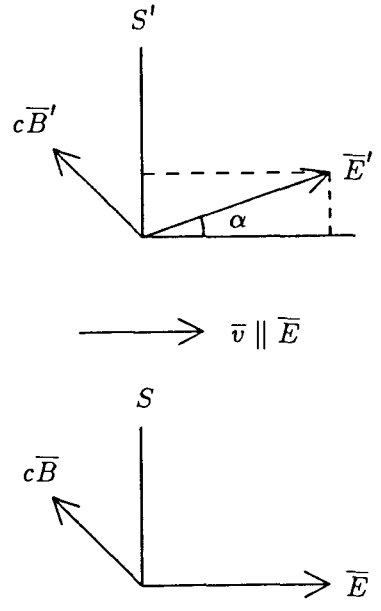


Figure 7.3.1 S' **moves in the direction of** $\overline{E}$.

The result is

$$\overline{E}' \cdot \overline{B}' = \overline{E} \cdot \overline{B} \tag{26}$$

According to (25) and (26), we can classify the electromagnetic field as follows:

b. Electromagnetic Field Classification

Free-Space Wave Fields: $\overline{E} \cdot \overline{B} = 0$ *and* $|\overline{E}| = |c\overline{B}|$

The field vectors $\overline{E}$ and $c\overline{B}$ are always perpendicular in direction and equal in magnitude. The magnitude changes from observer to observer, as is characteristic of a plane wave in free space. Consider the case $\overline{E} = \hat{x}E$ and $\overline{B} = \hat{y}B$ and assume that an observer S' is moving along the z axis of S. Then by (14*a*) and (14*b*) we immediately obtain

$$E' = E\left[\frac{1-\beta}{1+\beta}\right]^{1/2}$$

$$B' = B\left[\frac{1-\beta}{1+\beta}\right]^{1/2}$$

The amplitude of the wave field decreases as the velocity along the $\hat{z}$ direction increases. When it reaches c, namely, $\beta = 1$, the amplitude is zero. Thus an observer moving in the $\overline{E} \times \overline{B}$ direction at the velocity of light sees no field at all. An observer moving in a direction opposite to that of $\overline{E} \times \overline{B}$ at the velocity of light sees the amplitude approaching infinity. We can consider another simple case in which S' moves along the $\overline{E}$-field direction; the situation is shown in Figure 7.3.1. As the velocity becomes larger, the $\overline{E}'$-field vector tilts its direction so that $(-\overline{E}' \times c\overline{B}')$ tends to be parallel to $\overline{v}$. The angle of tilt of $\overline{E}'$ is $\alpha = \sin^{-1}\beta$, which increases as β increases.

Electric Fields: $\overline{E} \cdot \overline{B} = 0$ *and* $|\overline{E}| > |c\overline{B}|$

From $(14b)$, we see that there exists an observer moving in the $\overline{E} \times \overline{B}$ direction who experiences only an electric field and no magnetic field. The velocity of this observer is given by $\beta = |c\overline{B}| / |\overline{E}|$. Apparently $|\overline{v}|$ is smaller than c or $\beta < 1$. It is interesting to note that this observer, called S', is not the only one who does not experience a magnetic field. All observers moving along the $\overline{E}'$-field vector relative to S' also do not experience a magnetic field.

Magnetic Fields: $\overline{E} \cdot \overline{B} = 0$ *and* $|\overline{E}| < |c\overline{B}|$

This is the dual of the case above. Observer S' moving with a speed $\beta = |\overline{E}|/|c\overline{B}|$ along the $\overline{E} \times \overline{B}$ direction relative to S, and all other observers moving in the $\overline{B}'$ direction with respect to S' experience only a magnetic field and no electric field.

Wrench Fields: $\overline{E} \cdot \overline{B} \neq 0$

There are six cases in this class. It is clear that there are frames where $\overline{E}$ and $\overline{B}$ fields are parallel or antiparallel. When they are parallel, $\overline{E} \cdot \overline{B} > 0$, we have a positive wrench field. When they are antiparallel, $\overline{E} \cdot \overline{B} < 0$, we have a negative wrench field. The wrench field is electric if $|\overline{E}| > |c\overline{B}|$ because here the electric field dominates. The wrench field is magnetic if $|\overline{E}| < |c\overline{B}|$.

7.4 Transformation of Constitutive Relations

We have seen that, in the transformation of electromagnetic field vectors, $\overline{E}$ and $c\overline{B}$ transform together, forming the entity of quantity, which is a four-dimensional second-rank tensor in the Minkowski space.

The electromagnetic field vectors $\overline{H}$ and $c\overline{D}$ transform together and form the entity of excitation, which is also a four-dimensional second-rank tensor. Thus we write the constitutive relations in the $\overline{E}\,\overline{B}$ representation:

$$\begin{bmatrix} c\overline{D} \\ \overline{H} \end{bmatrix} = \overline{\overline{C}} \cdot \begin{bmatrix} \overline{E} \\ c\overline{B} \end{bmatrix} \tag{1}$$

where

$$\overline{\overline{C}} = \begin{bmatrix} \overline{\overline{P}} & \overline{\overline{L}} \\ \overline{\overline{M}} & \overline{\overline{Q}} \end{bmatrix} \tag{2}$$

is the constitutive matrix the elements of which are constitutive parameters. This representation provides a Lorentz covariant description of the constitutive relations.

The Lorentz transformation formulas for the electromagnetic field vectors can now be used to derive transformation laws for the constitutive relations. A medium at rest in one frame becomes a medium in motion when viewed from another frame. The derivation of equivalent constitutive relations for a moving medium in the laboratory frame is useful conceptually and practically. It is indeed true that a problem involving one moving medium can always be solved in its rest frame and the results can be transformed back to the laboratory frame. In practice, the Lorentz transformation method cannot be applied when more than two relatively moving media are involved because, in the rest frame of one medium, all others are in motion. Thus constitutive relations for moving media have to be determined.

a. Moving Isotropic Media

Consider two reference frames in relative uniform motion. In the reference frame S' there is an isotropic medium with permittivity ϵ' and permeability μ'. The constitutive matrix takes the form

$$\overline{\overline{C}}' = \begin{bmatrix} c\mu' \overline{\overline{I}} & \overline{\overline{0}} \\ \overline{\overline{0}} & \dfrac{1}{c\mu'} \overline{\overline{I}} \end{bmatrix} \tag{3}$$

In the laboratory frame S we obtain the constitutive matrix for the

moving isotropic medium

$$\begin{aligned}\overline{\overline{C}} &= \overline{\overline{L}}_6^{-1} \cdot \overline{\overline{C}}' \cdot \overline{\overline{L}}_6 \\ &= \frac{\gamma^2}{c'} \begin{bmatrix} (n^2-\beta^2)\overline{\overline{I}} - (n^2-1)\overline{\beta}\,\overline{\beta} & (n^2-1)\overline{\overline{\beta}} \\ (n^2-1)\overline{\overline{\beta}} & (1-n^2\beta^2\overline{\overline{I}}) + (n^2-1)\overline{\beta}\,\overline{\beta} \end{bmatrix}\end{aligned} \tag{4}$$

where $n^2 = c^2\mu'\epsilon'$ is the squared refractive index of the moving medium in its rest frame of reference. Clearly (4) reduces to (3) when $\beta = 0$. In a vacuum when $n = 1$, $\overline{\overline{C}}$ reduces to a constant $(1/c\mu_o)$ times a unit matrix. For an isotropic medium in motion, it becomes bianisotropic.

When the velocity is along the $\hat{z}$ direction, (1) becomes

$$\overline{\overline{C}} = \frac{\gamma^2}{c\mu'} \begin{bmatrix} n^2-\beta^2 & 0 & 0 & 0 & -(n^2-1)\beta & 0 \\ 0 & n^2-\beta^2 & 0 & (n^2-1)\beta & 0 & 0 \\ 0 & 0 & n^2(1-\beta^2) & 0 & 0 & 0 \\ 0 & -(n^2-1)\beta & 0 & 1-n^2\beta^2 & 0 & 0 \\ (n^2-1)\beta & 0 & 0 & 0 & 1-n^2\beta^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/\gamma^2 \end{bmatrix} \tag{5}$$

Note that, although we assume μ' and ϵ' to be scalar numbers, the derivation remains valid even if they are not scalar. We have to remember that μ' and ϵ' are measured in the rest frame of the medium. If they are dependent on parameters pertaining to their rest frame, these parameters must be properly transformed. For instance, when the medium in its rest frame is a plasma we need to transform the plasma frequency to the laboratory frame.

The constitutive relations for a moving isotropic medium can be written in the form of $\overline{B}$ and $\overline{D}$ expressed in terms of $\overline{E}$ and $\overline{H}$. The result is as follows:

$$\overline{B} = \mu'\overline{\overline{A}} \cdot \overline{H} - \overline{\Omega} \times \overline{E} \tag{6a}$$

$$\overline{D} = \epsilon'\overline{\overline{A}} \cdot \overline{E} + \overline{\Omega} \times \overline{H} \tag{6b}$$

where

$$\overline{\overline{A}} = \frac{1-\beta^2}{1-n^2\beta^2}\left[\overline{\overline{I}} - \frac{n^2-1}{1-\beta^2}\overline{\beta}\,\overline{\beta}\right] \tag{7}$$

and

$$\overline{\Omega} = \frac{n^2-1}{1-n^2\beta^2}\overline{\beta}/c \tag{8}$$

When $\bar{v}$ is in the $\hat{z}$ direction, $\overline{\overline{A}}$ becomes a diagonal matrix.

b. Moving Bianisotropic Media

We now consider the general case in which a bianisotropic medium with constitutive relations (1) is in motion. We assume that the velocity is in the $\hat{z}$ direction. This assumption for an unbounded medium does not impose a restriction on the validity of the results that are obtained because we can always rotate our coordinate system so that the z axis points in the direction of motion. This coordinate rotation certainly affects the elements of the constitutive matrix. But after the transformation, $\overline{\overline{C}}$ still takes the general form (3) except that the element values are different. By application of (7.3.19), the constitutive matrix is determined to be

$$\overline{\overline{P}} = \begin{bmatrix} \gamma^2\left[p_{xx} - \beta(l_{xy} - m_{yx}) - \beta^2 q_{yy}\right] & \gamma^2\left[p_{xy} - \beta(l_{xx} + m_{yy}) + \beta^2 q_{yx}\right] & \gamma(p_{xz} + \beta m_{yz}) \\ \gamma^2\left[p_{yx} + \beta(l_{yy} + m_{xx}) + \beta^2 q_{xy}\right] & \gamma^2\left[p_{yy} + \beta(l_{yx} - m_{yx}) - \beta^2 q_{xx}\right] & \gamma(p_{yz} - \beta m_{xz}) \\ \gamma(p_{zx} - \beta l_{zy}) & \gamma(p_{zy} + \beta l_{zx}) & p_{zz} \end{bmatrix} \tag{9a}$$

$$\overline{\overline{Q}} = \begin{bmatrix} \gamma^2\left[q_{xx} + \beta(m_{xy} - l_{yx}) - \beta^2 p_{yy}\right] & \gamma^2\left[q_{xy} - \beta(m_{xx} + l_{yy}) + \beta^2 p_{yx}\right] & \gamma(q_{xz} - \beta l_{yz}) \\ \gamma^2\left[q_{yx} + \beta(m_{yy} + l_{xx}) + \beta^2 p_{xy}\right] & \gamma^2\left[q_{yy} - \beta(m_{yx} - l_{xy}) - \beta^2 p_{xx}\right] & \gamma(q_{yz} + \beta l_{xz}) \\ \gamma(q_{zx} + \beta m_{zy}) & \gamma(q_{zy} - \beta m_{zx}) & q_{zz} \end{bmatrix} \tag{9b}$$

$$\overline{\overline{L}} = \begin{bmatrix} \gamma^2\left[l_{xx} + \beta(p_{xy} + q_{yx}) + \beta^2 m_{yy}\right] & \gamma^2\left[l_{xy} - \beta(p_{xx} - q_{yy}) - \beta^2 m_{yx}\right] & \gamma(l_{xz} + \beta q_{yz}) \\ \gamma^2\left[l_{yx} + \beta(p_{yy} - q_{xx}) - \beta^2 m_{xy}\right] & \gamma^2\left[l_{yy} - \beta(p_{yx} + q_{yx}) + \beta^2 q_{xx}\right] & \gamma(l_{yz} - \beta q_{xz}) \\ \gamma(l_{zx} + \beta p_{zy}) & \gamma(l_{zy} - \beta p_{zx}) & l_{zz} \end{bmatrix} \tag{9c}$$

$$\overline{\overline{M}} = \begin{bmatrix} \gamma^2\left[m_{xx} - \beta(q_{xy} + p_{yx}) + \beta^2 l_{yy}\right] \\ \gamma^2\left[m_{yx} - \beta(q_{yy} - p_{xx}) - \beta^2 l_{xy}\right] \\ \gamma(m_{zx} - \beta q_{zy}) \end{bmatrix}$$

$$\begin{bmatrix} \gamma^2\left[m_{xy}+\beta(q_{xx}-p_{yy})-\beta^2 l_{yx}\right] & \gamma(m_{xz}-\beta p_{yz}) \\ \gamma^2\left[m_{yy}+\beta(q_{yx}+p_{yx})+\beta^2 l_{xx}\right] & \gamma(m_{yz}+\beta p_{xz}) \\ \gamma(m_{zy}+\beta q_{zz}) & m_{zz} \end{bmatrix} \tag{9d}$$

It is noted that if a bianisotropic medium satisfies the symmetry conditions in its rest frame, then the moving bianisotropic medium also satisfies the symmetry conditions. We conclude that for moving nonabsorbing bianisotropic media, the 6×6 constitutive matrix $\overline{\overline{C}}$ can contain only up to 21 independent complex elements.

c. *Moving Gyrotropic Media*

Formulas (9) enable an observer to characterize a bianisotropic medium moving along the $\hat{z}$ direction. We observe that when a medium is in motion, it becomes bianisotropic. For a moving electrically gyrotropic medium with scalar permeability μ and permittivity tensor

$$\overline{\overline{\epsilon}} = \begin{bmatrix} \epsilon & i\epsilon_g & 0 \\ -i\epsilon_g & \epsilon & 0 \\ 0 & 0 & \epsilon_z \end{bmatrix} \tag{10}$$

we obtain

$$\overline{\overline{C}} = \frac{\gamma^2}{c\mu} \begin{bmatrix} n^2-\beta^2 & -n_g^2 & 0 & -\beta n_g^2 & -(n^2-1)\beta & 0 \\ -n_g^2 & n^2-\beta^2 & 0 & -(n^2-1)\beta & 0 & 0 \\ 0 & 0 & n_z^2 & 0 & 0 & 0 \\ \beta^2 n_g^2 & (n^2-1)\beta & 0 & 1-n^2\beta^2 & -\beta n_g^2 & 0 \\ (n^2-1)\beta & -\beta n_g^2 & 0 & -\beta n_g^2 & 1-n^2\beta^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{11}$$

where

$$n_g^2 = c^2\mu\epsilon_g \tag{12a}$$

$$n_z^2 = c^2\mu\epsilon_z \tag{12b}$$

The parameters that govern the gyrotropic nature of the medium must be carefully transformed. For example [Chawla and Unz, 1971], the plasma frequency is a Lorentz-invariant $\omega_p' = \omega_p$, while the cyclotron frequency transforms as $\omega_c' = \gamma\omega_c$. Remember that the static magnetic field is in the direction of motion. If the magnetic field were perpendicular to the direction of motion, we would have $\omega_c' = \gamma^2\omega_c$ instead. The applied frequency ω must also be properly transformed.

d. Moving Uniaxial Media

Constitutive relations for a moving uniaxial medium are also derived from the general formulas (9). We assume that in the rest frame of the moving medium the constitutive relations are as follows:

$$\bar{\bar{\epsilon}} = \begin{bmatrix} \epsilon' & 0 & 0 \\ 0 & \epsilon' & 0 \\ 0 & 0 & \epsilon'_z \end{bmatrix} \tag{13a}$$

$$\bar{\bar{\mu}} = \begin{bmatrix} \mu' & 0 & 0 \\ 0 & \mu' & 0 \\ 0 & 0 & \mu'_z \end{bmatrix} \tag{13b}$$

where the z axis coincides with the optic axis. Note that isotropic media and electric or magnetic uniaxial media are all special cases of (13). The primes indicate that these quantities are associated with the rest frame of the medium. In the laboratory frame, where the media appear to be moving uniformly with the velocity $\bar{v}$ along the $\hat{z}$ direction, the constitutive matrix of the medium is determined to be

$$\bar{\bar{C}} = \frac{1}{c\mu'} \begin{bmatrix} p & 0 & 0 & 0 & -l & 0 \\ 0 & p & 0 & l & 0 & 0 \\ 0 & 0 & p_z & 0 & 0 & 0 \\ 0 & -l & 0 & q & 0 & 0 \\ l & 0 & 0 & 0 & q & 0 \\ 0 & 0 & 0 & 0 & 0 & q_z \end{bmatrix} \tag{14}$$

Note that p, q, l, p_z, and q_z are all dimensionless quantities. The constitutive matrix can be transformed into the $\overline{E}\,\overline{H}$ and the $\overline{D}\,\overline{H}$ representations. We obtain

$$\bar{\bar{C}}_{EH} = \begin{bmatrix} \epsilon & 0 & 0 & 0 & \xi & 0 \\ 0 & \epsilon & 0 & -\xi & 0 & 0 \\ 0 & 0 & \epsilon_z & 0 & 0 & 0 \\ 0 & -\xi & 0 & \mu & 0 & 0 \\ \xi & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu_z \end{bmatrix} \tag{15}$$

$$\bar{\bar{C}}_{DB} = \begin{bmatrix} \kappa & 0 & 0 & 0 & \chi & 0 \\ 0 & \kappa & 0 & -\chi & 0 & 0 \\ 0 & 0 & \kappa_z & 0 & 0 & 0 \\ 0 & -\chi & 0 & \nu & 0 & 0 \\ \chi & 0 & 0 & 0 & \nu & 0 \\ 0 & 0 & 0 & 0 & 0 & \nu_z \end{bmatrix} \tag{16}$$

The element values of the constitutive matrices in different representations are summarized in Table 7.4.1. Remember that the velocity $\overline{v}$ is in the $\hat{z}$ direction. The matrix elements are velocity dependent. The primed quantities are measured in the rest frame of the media; if they are coordinate dependent, we must also transform them to the laboratory frame. For instance, in the case of a moving isotropic plasma, both the plasma frequency and the applied frequency must be transformed.

From Table 7.4.1 we see that the constitutive matrix $\overline{\overline{C}}$ becomes diagonal if $l = 0$ or $\mu'\epsilon' = 1/c^2\,(n = 1)$, as can sometimes be achieved in a dispersive medium. For an anisotropic plasma subject to a strong magnetic field in the $\hat{z}$ direction, $\epsilon' = \epsilon_o$. In such a medium we also have $\mu' = \mu_o$. Thus $l = 0, n = p = q = l$, and $\overline{\overline{C}}$ will be diagonal. This is an example of a moving medium that is not bianisotropic. All of the constitutive relations take the forms (14)–(16) whether the moving medium in its rest frame is an isotropic medium, a uniaxial crystal, a uniaxial gyrotropic plasma, or a magnetic and electric uniaxial medium.

e. Accelerated Media

The constitutive relations of an accelerated medium are also bianisotropic in form. They are space dependent and can be viewed as inhomogeneous. General formulations in arbitrary accelerating frames have been considered by some authors, and explicit forms for the constitutive relations have been proposed for rotating and linearly accelerated media. Not only relative motion but also absolute motion of both observer and medium play a crucial role in determining the constitutive relations.

7.5 Transformation of Frequency and Wave Vector

Consider a receiver and a transmitter in relative motion, and assume that the receiver receives a plane wave from the transmitter. According to the transmitter, the plane wave is described by

$$\begin{bmatrix} \overline{E}(r,t) \\ c\overline{B}(r,t) \end{bmatrix} = \begin{bmatrix} \overline{E}_0 \\ c\overline{B}_0 \end{bmatrix} \cos(\overline{k}\cdot\overline{r} - \omega t) \tag{1}$$

$\overline{E}\,\overline{B}$ *Representation*

$$p = \gamma(n^2 - \beta^2), \qquad p_z - an^2, \qquad a = \frac{\epsilon'_z}{\epsilon'}$$

$$q = \gamma^2(1 - n^2\beta^2), \quad q_z = \frac{1}{b}, \qquad b = \frac{\mu'_z}{\mu'},$$

$$l = \gamma^2\beta(n^2 - 1), \qquad qp + l^2 = n^2 = c^2\mu'\epsilon'$$

$\overline{E}\,\overline{H}$ *Representation*

$$\epsilon = \frac{(qp + l^2)}{c^2\mu' q} = \frac{\epsilon'(1 - \beta^2)}{(1 - n^2\beta^2)}, \quad \epsilon_z = \epsilon'_z$$

$$\mu = \frac{\mu'}{q} = \frac{\mu'(1 - \beta^2)}{(1 - n^2\beta^2)}, \qquad \mu_z = \mu'_z$$

$$\xi = \frac{-l}{cq} = \frac{-\beta(n^2 - 1)}{c(1 - n^2\beta^2)}$$

$\overline{D}\,\overline{B}$ *Representation*

$$\kappa = \frac{c^2\mu'}{q} = \frac{c^2\mu'(1 - \beta^2)}{(n^2 - \beta^2)}, \qquad \kappa_z = \frac{1}{\epsilon'_z}$$

$$\nu = \frac{(qp + l^2)}{p\mu'} = \frac{c^2\epsilon'(1 - \beta^2)}{(n^2 - \beta^2)}, \quad \nu_z = \frac{1}{\nu'_z}$$

$$\chi = \frac{cl}{p} = \frac{c\beta(n^2 - 1)}{(n^2 - \beta^2)}$$

Table 7.4.1 Constitutive parameters for moving media.

According to the receiver S', the plane wave is described by

$$\begin{bmatrix} \overline{E}'(\overline{r}', \overline{t}') \\ c\overline{B}'(\overline{r}', \overline{t}') \end{bmatrix} = \begin{bmatrix} \overline{E}'_0 \\ c\overline{B}'_0 \end{bmatrix} \cos(\overline{k}' \cdot \overline{r}' - \omega t') \tag{2}$$

Let the receiver S' move with uniform velocity $\overline{v}$ with respect to the transmitter S, and let primes denote quantities associated with S'.

According to the Lorentz transformation formulas, we have

$$\begin{bmatrix} \overline{E}_0 \\ c\overline{B}_0 \end{bmatrix} = \begin{bmatrix} \overline{\overline{\alpha}}^{-1} & -\overline{\overline{\beta}} \\ \overline{\overline{\beta}} & \overline{\overline{\alpha}}^{-1} \end{bmatrix} \cdot \begin{bmatrix} \overline{E}'_0 \\ c\overline{B}'_0 \end{bmatrix} \tag{3}$$

and

$$\overline{r} = \overline{\overline{\alpha}} \cdot \overline{r}' + \gamma \overline{\beta} ct' \tag{4a}$$

$$ct = \gamma(ct' + \overline{\beta} \cdot \overline{r}') \tag{4b}$$

The phase factor in (1) becomes

$$\overline{k} \cdot \overline{r} - \frac{\omega}{c} ct = (\overline{k} \cdot \overline{\overline{\alpha}} - \gamma \overline{\beta} \frac{\omega}{c}) \cdot \overline{r}' - \gamma \left[\frac{\omega}{c} - \overline{\beta} \cdot \overline{k} \right] ct'$$

Comparing with (2), this is equal to

$$\overline{k}' = \overline{\overline{\alpha}} \cdot \overline{k} - \gamma \overline{\beta} \frac{\omega}{c} \tag{5a}$$

$$\frac{\omega'}{c} = \gamma \left[\frac{\omega}{c} - \overline{\beta} \cdot \overline{k} \right] \tag{5b}$$

The transformation formula is identical to that for space and time coordinates with $\overline{r}$ replaced by $\overline{k}$ and t replaced by ω/c. With transformation formula (5), the phase of the plane wave in both frames is an invariant quantity. This invariance of phase, which enables us to deduce transformation formula (5), is referred to as the principle of phase invariance.

a. Aberration Effect

The aberration effect is a consequence of (5). The perpendicular component of $\overline{k}'$ is equal to that of $\overline{k}$, while the parallel component is changed by the motion. Consider an observer on Earth looking at a star at the zenith. Since the Earth is moving with respect to the star, a $\overline{k}'$ component antiparallel to $\overline{\beta}$ is generated. Thus the observer must tilt his telescope in the direction of the Earth's motion, just as, on a windless rainy day, a bicycle rider always tilts his umbrella in the forward direction. It is straightforward to determine from (5*a*) a relation for the angles between $\overline{\beta}$ and $\overline{k}'$ and between $\overline{\beta}$ and $\overline{k}$. Let θ denote the angle between $\overline{k}$ and $\overline{\beta}$, and θ' the angle between $\overline{k}'$

and $\overline{\beta}$. Recall that $\overline{\overline{\alpha}}$ is defined by $\overline{\overline{\alpha}} = \overline{\overline{I}} + (\gamma - 1)(\overline{\beta}\,\overline{\beta}/\beta^2)$. Cross and dot multiplying $(5a)$ by $\overline{\beta}$, we obtain

$$k' \sin\theta' = k \sin\theta$$
$$k' \cos\theta' = \gamma k \cos\theta - \frac{\gamma\beta}{c}$$

In an isotropic medium, $k = n\omega/c$, elimination of k' and k from these two equations gives

$$\tan\theta' = \frac{\tan\theta}{\gamma[1 - (\beta \sec\theta)/n]} \tag{6}$$

This is the relativistic formula for aberration.

b. Doppler Effect

The Doppler effect is a consequence of $(5b)$. Using the dispersion relation for isotropic media and letting the angle between $\overline{k}$ and $\overline{\beta}$ be θ, we find from $(5b)$ that

$$\omega' = \gamma\omega(1 - n\beta\cos\theta) \tag{7}$$

When the receiver is receiving from the transmitter, $\overline{\beta}$ and $\overline{k}$ are in the same direction. The frequency is shifted downward or red-shifted. When the receiver is approaching the transmitter, the frequency is shifted upward or blue-shifted. When the receiver is moving perpendicularly to $\overline{k}$, we have the transverse Doppler shift $\omega' = \gamma\omega$, which is a purely relativistic effect.

7.6 Plane Waves in Moving Uniaxial Media

We make use of the kDB system to study propagation of plane waves in unbounded moving uniaxial media. The constitutive relations in the $\overline{D}\,\overline{B}$ representation take the form

$$\overline{E} = \overline{\overline{\kappa}} \cdot \overline{D} + \overline{\overline{\chi}} \cdot \overline{B} \tag{1a}$$
$$\overline{H} = \overline{\overline{\gamma}} \cdot \overline{D} + \overline{\overline{\nu}} \cdot \overline{B} \tag{1b}$$

where

$$\overline{\overline{\kappa}} = \begin{bmatrix} \kappa & 0 & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & \kappa_z \end{bmatrix} \tag{2a}$$

$$\overline{\overline{\nu}} = \begin{bmatrix} \nu & 0 & 0 \\ 0 & \nu & 0 \\ 0 & 0 & \nu_z \end{bmatrix} \tag{2b}$$

$$\overline{\overline{\chi}} = \overline{\overline{\gamma}}^{+} = \begin{bmatrix} 0 & \chi & 0 \\ -\chi & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{2c}$$

The constitutive parameters κ, κ_z, ν, ν_z and χ are found in 7.4.

In the kDB system, the constitutive matrices become

$$\overline{\overline{\kappa}}_k = \begin{bmatrix} \kappa & 0 & 0 \\ 0 & \kappa\cos^2\theta + \kappa_z \sin^2\theta & (\kappa - \kappa_z)\sin\theta\cos\theta \\ 0 & (\kappa - \kappa_z)\sin\theta\cos\theta & \kappa\sin^2\theta + \kappa_z\cos^2\theta \end{bmatrix} \tag{3a}$$

$$\overline{\overline{\nu}}_k = \begin{bmatrix} \nu & 0 & 0 \\ 0 & \nu\cos^2\theta + \nu_z \sin^2\theta & (\nu - \nu_z)\sin\theta\cos\theta \\ 0 & (\nu - \nu_z)\sin\theta\cos\theta & \nu\sin^2\theta + \nu_z\cos^2\theta \end{bmatrix} \tag{3b}$$

$$\overline{\overline{\chi}}_k = \overline{\overline{\gamma}}_k^{+} = \begin{bmatrix} 0 & \chi\cos\theta & \chi\sin\theta \\ -\chi\cos\theta & 0 & 0 \\ -\chi\sin\theta & 0 & 0 \end{bmatrix} \tag{3c}$$

Substituting into (2.3.36), (2.3.37) and eliminating $\overline{B}$, we obtain the following equation for $\overline{D}$:

$$\begin{bmatrix} 1 - (\mu - \chi\cos\theta)^2/\kappa(\nu\cos^2\theta + \nu_z\sin^2\theta) & 0 \\ 0 & 1 - (\mu - \chi\cos\theta)^2/\kappa(\kappa\cos^2\theta + \kappa_z\sin^2\theta) \end{bmatrix} \cdot \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} = 0 \tag{4}$$

The phase velocities of the two characteristic waves are easily obtained. Other field components are found from (3b) and the constitutive relations. The results are listed in Table 7.6.1. Writing explicity in terms of components of the $\overline{k}$ vector and noting that $k_z^2 = k^2\cos^2\theta$ and $k_x^2 + k_y^2 = k^2\sin^2\theta$, we find the dispersion relation for the Type I wave

$$k_x^2 + k_y^2 + \frac{\nu k_z^2}{\nu_z} - \frac{(\omega - \chi k_z)^2}{\kappa\nu_z} = 0 \tag{5}$$

wave character \ modes	Type I wave	Type II wave
$\overline{D}_k$	$\begin{pmatrix}1\\0\\0\end{pmatrix}$	$\begin{pmatrix}0\\1\\0\end{pmatrix}$
$\overline{B}_k$	$\begin{pmatrix}0\\ \kappa/(u-\chi\cos\theta)\\0\end{pmatrix}$	$\begin{pmatrix}-(u-\chi\cos\theta)/\nu\\0\\0\end{pmatrix}$
$\overline{E}_k$	$\begin{pmatrix}\kappa u/(\nu-\chi\cos\theta)\\0\\0\end{pmatrix}$	$\begin{pmatrix}0\\(u-\chi\cos\theta)/\nu\\\left[\frac{\chi}{\nu}(u-\chi\cos\theta)+(\kappa_z-\kappa)\cos\theta\right]\sin\theta\end{pmatrix}$
$\overline{H}_k$	$\begin{pmatrix}0\\u\\-\chi+\frac{\kappa(\nu_z-\nu)\cos\theta}{u-\chi\cos\theta}\sin\theta\end{pmatrix}$	$\begin{pmatrix}-u\\0\\0\end{pmatrix}$
u	$\chi\cos\theta\pm\sqrt{\kappa(\nu\cos^2\theta+\nu_z\sin^2\theta)}$	$\chi\cos\theta\pm\sqrt{\nu(\kappa\cos^2\theta+\kappa_z\sin^2\theta)}$
Dispersion Relation	$k_x^2+k_y^2+\frac{\nu}{\nu_z}k_z^2-\frac{1}{\kappa\nu_z}(\omega-\chi k_z)^2=0$	$k_x^2+k_y^2+\frac{\kappa}{\kappa_z}k_z^2-\frac{1}{\nu\kappa_z}(\omega-\chi k_z)^2=0$

Table 7.6.1 Characteristic waves in moving uniaxial medium.

which propagates with velocity

$$u=\chi\cos\theta\pm\sqrt{\kappa(\nu\cos^2\theta+\nu_z\sin^2\theta)} \tag{6}$$

The dispersion relation for the Type II wave is

$$k_x^2+k_y^2+\frac{\kappa k_z^2}{\kappa_z}-\frac{(\omega-\chi\kappa_z)^2}{\nu k_z}=0 \tag{7}$$

which propagates with velocity

$$u=\chi\cos\theta\pm\sqrt{\nu(\kappa\cos^2\theta+\kappa_z\sin^2\theta)} \tag{8}$$

It is interesting to note the $\pm$ signs in (6) and (8). The plus sign corresponds to waves propagating in the direction of medium motion; the negative sign, to waves propagating opposite to the direction of medium motion. For the same type of wave, the magnitude of the velocity in opposite directions are now different, as opposed to all previous cases, in which they were the same. In Figure 7.6.1, we plot the

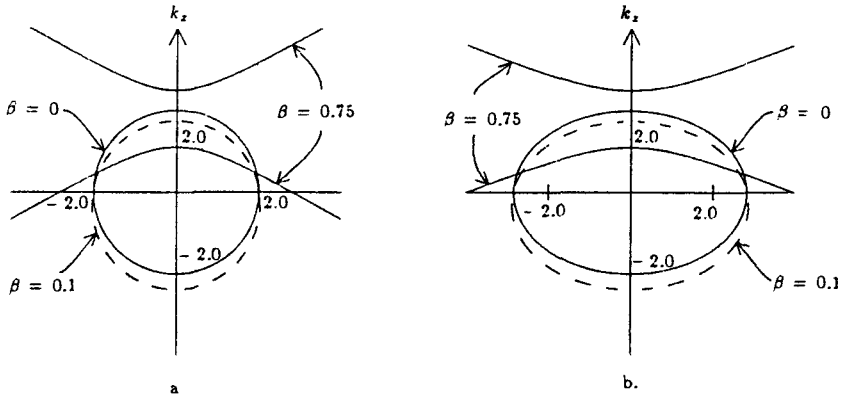


Figure 7.6.1 k **surfaces for moving isotropic medium.**

k surfaces for a moving isotropic medium with $n = 1$ in its rest frame and for a moving uniaxial medium with $n = 1$ and $a = b = 2$. For $1 - n^2\beta^2 > 0$, the surface is an ellipse rotated about the k_z axis; this is the non-relativistic case. For $1 - n^2\beta^2 < 0$, the k surface becomes a hyperbola rotated about the k_z axis; this is the relativistic case. We call this high-velocity region the Čerenkov zone. The velocity that separates the non-relativistic zone and the Čerenkov zone is $\beta = \pm 1/n$, which is equal to the velocity of light in the rest frame of the moving medium.

To facilitate further discussion, we make use of Table 7.6.1 and write (5) and (7) explicitly in terms of β dependence. After some manipulations, we obtain

$$k_x^2 + k_y^2 + b\frac{1 - n^2\beta^2}{1 - \beta^2}\left[k_z - \frac{n + \beta}{n\beta + 1}\frac{\omega}{c}\right] \cdot \left[k_z - \frac{n - \beta}{n\beta - 1}\frac{\omega}{c}\right] = 0 \quad (9)$$

$$k_x^2 + k_y^2 + a\frac{1 - n^2\beta^2}{1 - \beta^2}\left[k_z - \frac{n + \beta}{n\beta + 1}\frac{\omega}{c}\right] \cdot \left[k_z - \frac{n - \beta}{n\beta - 1}\frac{\omega}{c}\right] = 0 \quad (10)$$

We examine two cases. First, consider a wave propagating in the $\hat{x}$ direction perpendicular to the medium velocity, $k_z = 0$. The $\overline{k}$ vectors become

$$\overline{k} = \hat{x}k_x = \pm\hat{x}\frac{\omega}{c}\left[b\frac{n^2 - \beta^2}{1 - \beta^2}\right]^{1/2} \quad (11)$$

$$\overline{k} = \hat{x}k_x = \pm\hat{x}\frac{\omega}{c}\left[a\frac{n^2-\beta^2}{1-\beta^2}\right]^{1/2} \tag{12}$$

The $\pm$ sign distinguishes waves propagating in the positive and the negative $\hat{x}$ directions. As β increase from 0 to 1, k increases from $an\omega/c$ or from $bn\omega/c$ to infinity. Thus the velocity along the $\hat{z}$ direction is zero when the medium velocity approaches the velocity of light in vacuum.

Second, consider a wave propagating in the direction of medium motion, $k_x = k_y = 0$. The two types of waves degenerate into one, and the k vectors become

$$\overline{k} = \hat{z}\frac{n+\beta}{n\beta+1}\frac{\omega}{c} \tag{13}$$

$$\overline{k} = \hat{z}\frac{n-\beta}{n\beta-1}\frac{\omega}{c} \tag{14}$$

Equation (13) corresponds to waves propagating in the positive $\hat{z}$ direction, and (14) to waves propagating in the negative $\hat{z}$ direction. For the wave propagating in the positive $\hat{z}$ direction, we observe that, as β increases from 0 to 1, k decreases from $n\omega/c$ to ω/c. The corresponding velocity of the wave increases from c/n to c. For the wave propagating in the negative $\hat{z}$ direction, we observe that, as β increases from 0 to $1/n$, k changes from $-n\omega/c$ to $-\infty$, and the velocity changes from $-c/n$ to 0. As β further increases from $1/n$ to 1, k reverses sign and decreases from infinity to ω/c. In the Čerenkov zone the negatively propagating wave now propagates in the positive $\hat{z}$ direction. As β approaches 1, the velocity approaches c. In all cases, the wave appears to be dragged by the motion of the medium. This phenomenon is referred to as the *Fizeau-Fresnel drag.*

7.7 Phase Matching at Moving Boundaries

Previously, we have studied phase-matching conditions for a stationary boundary surface. When the boundary is moving, we let its velocity be

$$\overline{v} = \hat{x}v_x + \hat{v}_y + \hat{z}v_z \tag{1}$$

At $t = 0$, the surface is at $x = 0$. At other times, the boundary surface is at $x = v_x t$. The incident, the reflected, and the transmitted waves

now have space–time dependence as follows:

$$\begin{aligned} \text{incident}: &\quad e^{ik_x x+ik_y y+ik_z z-i\omega t} \\ \text{reflected}: &\quad e^{ik_{rx} x+ik_{ry} y+ik_{rz} z-i\omega_r t} \\ \text{transmitted}: &\quad e^{ik_{tx} x+ik_{ty} y+ik_{tz} z-i\omega_t t} \end{aligned}$$

where we distinguish between the angular frequencies ω, ω_r and ω_t. We require that the moving boundary conditions (1.5.19)–(1.5.22) be satisfied, which gives

$$\begin{aligned} &k_x(v_x t) + k_y(y + v_y t) + k_z(z - v_z t) - \omega t \\ &\quad = k_x(v_x t) + k_{ry}(t + v_y t) + k_{rz}(z + v_z t) - \omega_r t \\ &\quad = k_{tz}(v_x t) + k_{ty}(y + v_y t) + k_{tz}(z + v_z t) - \omega_t t \end{aligned} \tag{2}$$

Since these equalities must hold for all y, z and t, we conclude that

$$k_y = k_{ry} = k_{ty} \tag{3}$$

$$k_z = k_{rz} = k_{tz} \tag{4}$$

$$k_x v_x - \omega = k_{rx} v_x - \omega_r = k_{tx} v_x - \omega_t \tag{5}$$

Thus the phase-matching conditions are the same as for the case of a stationary boundary; tangential components of the wave vectors are continuous. The frequencies of the three waves are now different; they are determined only by the normal component of the velocity.

Consider the simple case of a wave normally incident from an isotropic medium upon another isotropic medium moving in the $\hat{x}$ direction toward the wave. The incident wave vector is

$$\overline{k} = \hat{x} k_x = -\hat{x} n \frac{\omega}{c} \tag{6}$$

and the reflected vector is

$$\overline{k}_r = \hat{x} k_{rx} = \hat{x} n \frac{\omega_r}{c} \tag{7}$$

Substituting (6)–(7) into (5) we find

$$\omega_r = \omega \frac{1 + n\beta}{1 - n\beta} \tag{8}$$

$$k_r = -k \frac{1 + n\beta}{1 - n\beta} = n \frac{\omega}{c} \frac{1 + n\beta}{1 - n\beta} \tag{9}$$

where $\beta = v_x/c$. The reflected wave is seen to be Doppler-shifted toward the high-frequency side. Its wavenumber is also increased by the same amount. In order to determine the transmitted wavenumber and angular frequencies, we must know the dispersion relations for a moving medium.

7.8 Guided Waves in a Moving Dielectric Slab

Consider TE modes guided by a slab medium moving between two identical stationary isotropic media. Let the guidance direction be parallel to the direction of motion of the slab. We follow the graphical approach to determine the cutoff wavenumbers and the propagation constant k_z.

The constitutive parameters for a moving medium has been determined in 7.4. We use subscript 1 to denote the parameters pertaining to the moving dielectric slab waveguide. The transverse wavenumber k_{1x} inside the waveguide, by making use of the dispersion relations, may be cast into the following form:

$$k_{1x}^2 = p_1 k_o^2 - 2l_1 k_z k_o - q_1 k_z^2 \tag{1}$$

where $k_o = \omega/c$ and $k_z = k_o(n^2 + \alpha_x^2/k_o^2)^{1/2}$. Equation (3.5.45) which follows from the guidance condition, remains unchanged. However, (3.5.46) now becomes

$$\begin{aligned}(\alpha_x d)^2 + (k_{1x}d)^2 &= (k_x^2 - k^2)d^2 + (pk_o^2 - 2lk_o k_z - qk_z^2)d^2 \\ &= (k_o d)^2 \left\{ \frac{n_1^2 - \beta^2}{1-\beta^2} - n^2 - 2\beta \frac{n_1^2 - 1}{1-\beta^2} \sqrt{n^2 + \left(\frac{\alpha_x}{k_o}\right)^2} \right. \\ &\qquad \left. + \beta^2 \frac{n_1^2 - 1}{1-\beta^2} \left[n^2 + \left(\frac{\alpha_x}{k_o}\right)^2 \right] \right\}\end{aligned} \tag{2}$$

The shape of the curve described by this equation is a function of the slab velocity. At $\beta = 0$, the curve is a circular arc, the expected result for a stationary slab. When the isotropic medium surrounding the slab is free space, $n = 1$ and this equation becomes

$$(\alpha_x d)^2 + (k_{1x}d)^2 = (k_o d)^2 \left[\frac{n_1^2 - 1}{1-\beta^2} \right] \left[1 - \beta \sqrt{1 + \left(\frac{\alpha_x}{k_o}\right)^2} \right]^2 \tag{3}$$

a result that can be derived by applying the Lorentz transformation directly. We observe that at cutoff $k_o d = \sqrt{(1+\beta)/(1-\beta)}m\pi\sqrt{n_1^2-1}$; the cutoff wavenumber increases as β increases. In general, we see that as β increases, α_x decreases, and so does k_x. Thus the guided waves possess higher cutoff frequencies and propagate at larger phase velocities. A parallel analysis results in similar conclusions for the TM waves.

Next, we consider an isotropic medium moving between two perfectly conducting parallel plates. The guidance condition for both TE and TM modes is

$$\frac{m\pi}{d} = k_{1x} = \gamma\sqrt{(n_1^2-\beta^2)k_o^2 - 2\beta(n_1^2-1)k_o k_z - (1-n_1^2\beta^2)k_z^2} \tag{4}$$

The second equality follows from (1). Note that $m\pi/d$ is the cutoff wavenumber for the m th mode when the medium is stationary. Solving for k_z, we obtain from this equation

$$k_z = \frac{k_o}{1-n_1^2\beta^2}\left\{-\beta(n_1^2-1) \pm (1-\beta^2)\cdot\sqrt{n_1^2 - \left(\frac{1-n_1^2\beta^2}{1-\beta^2}\right)\left(\frac{m\pi}{k_o d}\right)^2}\right\} \tag{5}$$

Observe that cutoff occurs when k_z becomes imaginary. When the velocity of the medium exceeds the Čerenkov velocity so that $n_1\beta > 1$, k_z will always be real for real n_1 and no cutoff will occur. The propagation constant k_z is always positive when

$$\beta(n_1^2-1) \geq \left[n_1^2(1-\beta^2)^2 - (1-\beta^2)^2(1-n_1^2\beta^2)(m\pi/k_o d)\right]^{1/2}$$

or, equivalently, $k_o d \geq m\pi\left[m\pi(1-\beta^2)/(n_1^2-\beta^2)\right]^{1/2}$.

In the low-velocity regime when $n_1\beta < 1$, cutoff occurs for $k_o d \leq m\pi(1-n_1^2\beta^2)/(1-\beta^2)n_1^2$. Comparing this with the stationary case in which $k_c d = m\pi/n_1^2$, we see that motion of the medium always lowers the cutoff wavenumber. For frequencies above cutoff but with the square-root term smaller than the first term in the equation, the phase velocities of the guided waves are all in the negative $\hat{z}$ direction. Phase velocities in both directions become possible when $k_o d \geq m\pi\left[(1-\beta^2)/(n_1^2-\beta^2)\right]^{1/2}$.

It is of interest to investigate the power flow carried by each mode in the waveguide. Consider the TE modes with

$$E_y = E_m \sin \frac{m\pi x}{d} e^{ik_z z}$$

The magnetic field components are determined from Maxwell's equations and the constitutive relations for the moving medium:

$$\begin{aligned}\overline{B} &= \frac{1}{i\omega}\nabla \times \overline{E} \\ &= \left[-\hat{x}\frac{k_z}{\omega}E_m \sin \frac{m\pi x}{d} - \hat{z}i\frac{m\pi}{\omega d}E_m \cos \frac{m\pi x}{d}\right] e^{ik_z z}\end{aligned}$$

$$\begin{aligned}\overline{H} &= \frac{1}{c\mu'}\left[\hat{x}(-lE_y + qc\beta_x) + \hat{z}(c\beta_z)\right] \\ &= \hat{x}\frac{1}{c\mu'}\left[-l - \frac{qck_z}{\omega}\right] E_y - \hat{z}i\frac{m\pi}{\omega\mu' d}E_m \cos \frac{m\pi x}{d} e^{ik_z z}\end{aligned}$$

The power flow in the $\hat{z}$ direction after integrating over the waveguide cross section is found to be

$$\begin{aligned}P_z &= -\int_0^d dx \frac{1}{2} Re(E_y H_x^*) \\ &= Re \sum_{m=1}^{\infty} \frac{d}{2c\mu'}\left[l + \frac{qk_z}{k_o}\right]^* |E_m|^2 \\ &= Re \sum_{m=1}^{\infty} \frac{d}{2c\mu'}\left[\pm\sqrt{n_1^2 - \left[\frac{1 - n_1^2\beta^2}{1 - \beta^2}\right]\left[\frac{m\pi}{k_o d}\right]^2}\right]^* |E_m|^2\end{aligned}$$

Thus each individual mode carries its own power; the total power is the sum of all individual components. At $\beta = 0$ this result reduces to the case of a stationary medium. Above the Čerenkov velocity, the square root always gives real values and all modes carry time-average power. Below the Čerenkov velocity, modes below cutoff will not carry time-average power. In all cases, the $\pm$ sign in front of the square root indicates that power can propagate in both positive and negative $\hat{z}$ directions, as opposed to the phase velocities, which in some cases can be in only one direction. In this velocity range, guided backward waves can be generated in the moving medium.

7.9 Guided Waves in Moving Gyrotropic Media

For guided waves in anisotropic and bianisotropic media the wave equations for E_z and H_z are usually coupled. We illustrate this with the general case of a bianisotropic medium realized by a moving gyrotropic medium. The gyrotropic medium in its rest frame has the tensor permittivity

$$\bar{\bar{\epsilon}}' = \begin{bmatrix} \epsilon' & -i\epsilon'_g & 0 \\ i\epsilon'_g & \epsilon' & 0 \\ 0 & 0 & \epsilon'_z \end{bmatrix} \tag{1}$$

In the laboratory frame the constitutive matrix is given by (7.4.11). We transform to $\overline{E}\,\overline{H}$ representation and obtain the following constitutive relations:

$$\overline{D} = \bar{\bar{\epsilon}}_s \cdot \overline{E}_s + \epsilon'_z \overline{E}_z + \bar{\bar{\xi}}_s \cdot \overline{H}_s \tag{2a}$$

$$\overline{B} = \bar{\bar{\mu}}_s \cdot \overline{E}_s + \mu'_z \overline{E}_z - \bar{\bar{\xi}}_s \cdot \overline{H}_s \tag{2b}$$

where

$$\bar{\bar{\epsilon}}_s = \epsilon' \begin{bmatrix} a & -ia_g \\ ia_g & a \end{bmatrix} = \frac{(1-\beta^2)\epsilon^2}{n^2\left[(1-n^2\beta^2)^2 - n_g^4\beta^4\right]} \cdot \begin{bmatrix} n^2(1-n^2\beta^2) + n_g^4\beta^2 & -in_g^2 \\ in_g^2 & n^2(1-n^2\beta^2) + n_g^4\beta^2 \end{bmatrix} \tag{3}$$

$$\bar{\bar{\mu}}_s = \mu' \begin{bmatrix} b & -ib_g \\ ib_g & b \end{bmatrix} = \frac{(1-\beta^2)\mu'}{(1-n^2\beta^2)^2 - n_g^4\beta^4} \begin{bmatrix} 1-n^2\beta^2 & -in_g^2\beta^2 \\ in_g^2\beta^2 & 1-n^2\beta^2 \end{bmatrix} \tag{4}$$

$$\bar{\bar{\xi}}_s = \frac{1}{c} \begin{bmatrix} -i\xi_g & -\xi \\ \xi & -i\xi_g \end{bmatrix} = \frac{(1-\beta^2)/c}{(1-n^2\beta^2)^2 - n_g^4\beta^4} \cdot \begin{bmatrix} -n_g^2\beta & -\gamma^2\beta\left[(n^2-1)(1-n^2\beta^2)+n_g^4\beta^2\right] \\ \gamma^2\beta\left[(n^2-1)(1-n^2\beta^2)+n_g^4\beta^2\right] & -in_g^2\beta \end{bmatrix} \tag{5}$$

At $\beta = 0$, (3) reduces to (1), $\bar{\bar{\mu}}_s = \mu'\bar{\bar{I}}$, $\bar{\bar{\xi}}_s = 0$, and we have the gyrotropic medium at rest.

We can express the transverse components in terms of the longitudinal components E_z and H_z and derive wave equations for E_z and

H_z. We find

$$\overline{\nabla}_s \times \overline{E}_z = i\omega\overline{\overline{\mu}}_s \cdot \overline{H}_s \cdot \overline{\overline{d}} \cdot \overline{E}_s \tag{6a}$$

$$\overline{\nabla}_s \times \overline{H}_z = -i\omega\overline{\overline{\epsilon}}_s \cdot \overline{H}_s - \overline{\overline{d}} \cdot \overline{H}_s \tag{6b}$$

$$\overline{\nabla}_s \times \overline{E}_s = i\omega\mu'_z \overline{H}_z \tag{7a}$$

$$\overline{\nabla}_s \times \overline{H}_s = -i\omega\epsilon'_z \overline{E}_z \tag{7b}$$

where

$$\overline{\overline{d}} = \begin{bmatrix} d_g & -id \\ id & d_g \end{bmatrix} = \begin{bmatrix} \omega\xi_g/c & -i(k_z + \omega\xi/c) \\ i(k_z + \omega\xi/c) & \omega\xi_g/c \end{bmatrix} \tag{8}$$

In terms of E_z and H_z, the transverse components are

$$\overline{E}_s = (\overline{\overline{I}} - \omega^2\overline{\overline{d}}^{-1} \cdot \overline{\overline{\mu}}_s \cdot \overline{\overline{d}}^{-1} \cdot \overline{\overline{\epsilon}}_s)^{-1} \cdot \left[-\overline{\overline{d}}^{-1} \cdot (\overline{\nabla}_s \times \overline{E}_z) - i\omega\overline{\overline{d}}^{-1} \cdot \overline{\overline{\mu}}_s \cdot \overline{\overline{d}}^{-1} (\overline{\nabla} \times \overline{H}_z) \right] \tag{9a}$$

$$\overline{H}_s = (\overline{\overline{I}} - \omega^2\overline{\overline{d}}^{-1} \cdot \overline{\overline{\epsilon}}_s \cdot \overline{\overline{d}}^{-1} \cdot \overline{\overline{\mu}}_s)^{-1} \cdot \left[-\overline{\overline{d}}^{-1} \cdot (\overline{\nabla}_s \times \overline{H}_z) - i\omega\overline{\overline{d}}^{-1} \cdot \overline{\overline{\epsilon}}_s \cdot \overline{\overline{d}}^{-1} (\overline{\nabla} \times \overline{E}_z) \right] \tag{9b}$$

after considerable algebraic manipulations, the wave equations for the longitudinal field components are determined to be

$$\left[\nabla_s^2 + \frac{\epsilon'_z}{\epsilon'} k^2 e \right] E_z = i\omega\mu' h_g H_z \tag{10a}$$

$$\left[\nabla_s^2 + k^2 h \right] H_z = -i\omega\epsilon'_z e_g E_z \tag{10b}$$

where

$$e = \frac{1}{b}\left[b^2 - b_g^2 + \frac{(bd - b_g d_g)^2}{d_g^2 - k^2 ab} \right] \tag{11a}$$

$$h = \frac{1}{a}\left[a^2 - a_g^2 + \frac{(ad - a_g d_g)^2}{d_g^2 - k^2 ab} \right] \tag{11b}$$

$$e_g = \frac{1}{a}\left[ab_g - a_g d - \frac{(ad - a_g d_g)(dd_g - kab_g)}{d_g^2 - k^2 ab} \right] \tag{11c}$$

$$h_g = \frac{1}{b}\left[bb_g - b_g d - \frac{(bd - b_g d_g)(dd_g - ka_g b)}{d_g^2 - k^2 ab} \right] \tag{11d}$$

The two equations in (10) for E_z and H_z are coupled. Thus the guided wave modes are hybrid because the wave equations are coupled. In the stationary case when $\beta = 0$, the wave equations are decoupled and the hybrid modes are consequences of boundary conditions.

Obtaining solutions to the coupled equations in (10) can be facilitated by transforming them into decoupled homogeneous Helmholtz equations. We define

$$\psi_j = E_z - i\alpha_j H_z \qquad j = 1, 2 \tag{12}$$

then multiply (10*b*) by i_j, and subtract the result from (10*a*). Using (12) to eliminate E_z and requiring that the coefficient for H_z be zero, we obtain the following second-order equation for α_j:

$$\omega\epsilon_z e_j \alpha_j^2 - k^2 \left[h - \frac{\epsilon_z'}{\epsilon'} e\right] \alpha_j - \omega\mu h_g = 0 \tag{13}$$

The two roots for α_j from (13) are the values for α_1 and α_2. With (13) satisfied, (10*a*) and (10*b*) combine to yield a single second-order, two-dimensional, scalar, homogeneous Helmholtz equation:

$$\nabla_s^2 \psi_j + q_j^2 \psi_j = 0 \tag{14}$$

where

$$\begin{aligned} q_j &= \frac{\epsilon_z'}{\epsilon'} k^2 e + \alpha_j \omega \epsilon_z' e_g \\ &= k^2 h + \frac{1}{\alpha_j} \omega\mu h_g \end{aligned} \tag{15}$$

For waveguides of conventional cross section (rectangular or circular), ψ_j can be determined from (14) in an appropriate coordinate system. With ψ_1 and ψ_2 known, we have from (12),

$$E_z = \frac{1}{\alpha_1 - \alpha_2}(\alpha_2 \psi_1 - \alpha_1 \psi_2) \tag{16a}$$

$$H_z = -\frac{1}{\alpha_1 - \alpha_2}(\psi_1 - \psi_2) \tag{16b}$$

The transverse field components can then be derived from (9) and be made to satisfy the boundary conditions. It is obvious that for guided

waves in a moving gyrotropic medium the modes are hybrid. Even when the gyrotropic medium is stationary, the two wave equations for E_z and H_z are still coupled and the modes are hybrid. The two wave equations will be decoupled if the medium is uniaxial, regardless of whether it is stationary or in motion, for then $\epsilon_g = 0$ and all parameters with subscript g will vanish.

7.10 Four-Dimensional Notations

a. Contravariant and Covariant Vectors

One can imagine a four-dimensional (4D) space composed of coordinates formed by time and three-dimensional space. Time–space coordinates of a physical event possess properties of a vector in 4D space. Let us denote the four components of an event by

$$x^0 = ct, \quad x^1 = x, \quad x^2 = y, \quad x^3 = z, \tag{1}$$

where we use the superscript 0 to denote the time component, and 1, 2, 3 to denote the space components. A popular convention in special relativity is to write the time component with the subscript 4 and designate it as imaginary. This should be carefully distinguished from the imaginary notations used in quantum theory and in wave theory. The notation that we use does not require an imaginary signature, but we must distinguish between superscripts and subscripts. This notation is readily generalized when general relativity is considered.

The transformation of the time–space coordinate vector from one observer to another is given by the Lorentz transformations. Under the Lorentz transformation,

$$x^2 + y^2 + z^2 - c^2t^2 = x'^2 + y'^2 + z'^2 - c^2t'^2 \tag{2}$$

is an invariant quantity independent of velocity. The square root of (2) expresses the magnitude of a 4D vector and in effect defines the transformation. As numerical values of other physical quantities change from one frame to another, this number stays unchanged in all frames. We note that in 4D space the magnitude of a vector can now be imaginary as well as real. A 4D vector is called a *spacelike* vector, a *null* vector, or a *timelike* vector, according to whether its magnitude is real, zero,

or imaginary, respectively. The time–space coordinates of two physical events (ct_1, x_1, y_1, z_1) and (ct_2, x_2, y_2, z_2) form a four-vector with magnitude squared

$$(X_1 - X_2)^2 = -c^2(t_1 - t_2)^2 + (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2$$

which is a Lorentz invariant. When the vector is timelike, $(X_1 - X_2)^2$ is negative. We can always find a moving observer such that, in his rest frame, the two events occur at the same location but at different times. He therefore observes the two events in person at times t_1' and t_2' :

$$-c(t_1' - t_2')^2 = (X_1 - X_2)^2$$

When the vector is a null vector, $|r_1 - r_2|^2 - c^2(t_1 - t_2)^2 = 0$ and an observer has to move at velocity c in order to see the two events in person. When the vector is spacelike, there exists an observer in whose frame the two events occur at different locations but simultaneously. The 4D space with coordinates governed by the Lorentz transformation laws is called *Minkowski space.*

To visualize the fourth dimension, imagine that this 4D space is spanned by four unit base vectors,

$$\hat{e}_\alpha = (\hat{e}_0, \hat{e}_1, \hat{e}_2, \hat{e}_3) \tag{3}$$

Any 4D vector X can thus be expressed in terms of $\hat{e}_\alpha$:

$$X = x^\alpha \hat{e}_\alpha = x^0\hat{e}_0 + x^1\hat{e}_1 + x^2\hat{e}_2 + x^3\hat{e}_3 \tag{4}$$

In (4) we use the Einstein summation convention: the repeated Greek index α implies summation from 0 to 3. We shall use Greek letters to indicate 0 to 3, and Roman letters to denote 1 to 3. The square of the length of X is defined as the scalar product of X with itself:

$$X^2 = \hat{e}_\alpha \cdot \hat{e}_\beta \, x^\alpha x^\beta \tag{5}$$

In view of (2),

$$X^2 = x^2 + y^2 + z^2 - c^2t^2 \tag{6}$$

Thus we must have

$$\hat{e}_\alpha \cdot \hat{e}_\beta = 0 \qquad \text{for} \quad \alpha \neq \beta \tag{7a}$$

$$\hat{e}_i \cdot \hat{e}_i = 1 \qquad \text{for} \quad i = 1,2,3 \tag{7b}$$

$$\hat{e}_0 \cdot \hat{e}_0 = -1 \tag{7c}$$

By (7*a*), all four base vectors are orthogonal to one another; by (7*b*), the three-space base vectors have unit magnitude as usual; and, by (7*c*), the zeroth (or fourth) base vector describing the fourth dimension possesses a magnitude squared of -1. It follows that the length of the zeroth base vector is imaginary. These four base vectors may be called *contravariant base* vectors which span a contravariant 4D space. A vector expressed in terms of the contravariant base vectors is called a *contravariant* vector.

We can define a set of covariant base vectors $\hat{e}^0$, $\hat{e}^1$, $\hat{e}^2$, and $\hat{e}^3$ such that

$$\hat{e}^0 = -\hat{e}_0 \tag{8a}$$

$$\hat{e}^i = \hat{e}_i \qquad i = 1,2,3 \tag{8b}$$

The product of $\hat{e}^0 \cdot \hat{e}_0 = -\hat{e}_0 \cdot \hat{e}_0 = 1$ in view of (7*c*). The vector $\hat{e}^0$ also has a magnitude squared of -1 as $\hat{e}^0 \cdot \hat{e}^0 = \hat{e}_0 \cdot \hat{e}_0$ by definitions of (8*a*) and (7*c*). These four base vectors $\hat{e}^\alpha$ may be called covariant base vectors which describe a covariant 4D space. A vector expressed in terms of the covariant base vectors is called a *covariant* vector. We write

$$X = x_\alpha \hat{e}^\alpha \tag{9}$$

Components of X in the new base are now denoted as x_α, which, in view of (8), are related to x^α by

$$x^0 = -x_0 \tag{10a}$$

$$x^i = x_i \tag{10b}$$

The contravariant components of X are denoted by superscripts, and its covariant components by subscripts. We define

$$\eta_{\alpha\beta} = \hat{e}_\alpha \hat{e}_\beta = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{11a}$$

and

$$\eta^{\alpha\beta} = \hat{e}^\alpha \hat{e}^\beta = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{11b}$$

to express the transformation between the two sets of base vectors and the contravariant and covariant components of a vector:

$$\hat{e}^{\alpha} = \eta^{\alpha\beta}\hat{e}_{\beta}, \qquad \hat{e}_{\alpha} = \eta_{\alpha\beta}\hat{e}^{\beta} \tag{12a}$$

$$\hat{x}^{\alpha} = \eta^{\alpha\beta}\hat{x}_{\beta}, \qquad \hat{x}_{\alpha} = \eta_{\alpha\beta}\hat{x}^{\beta} \tag{12b}$$

Equation (12) is equivalent to (8) and (10). The scalar product of two vectors is defined to be the summation over the contravariant components of one vector and the corresponding covariant components of another. Thus the magnitude squared of x^{α} is

$$x^2 = x^{\alpha}x_{\alpha} = \eta_{\alpha\beta}x^{\alpha}x^{\beta} \tag{13}$$

in view of (12*b*).

We summarize the rules for using the indices notation as follows:

(i) When the index of a vector is raised or lowered, the zeroth component of the vector changes sign, while the other components remain unchanged.

(ii) When an index is denoted by a Greek letter, it ranges from 0 to 3. When an index is denoted by a Roman letter, it ranges from 1 to 3.

(iii) When an index is repeated on the same side of an equation, a summation over the index is implied. Summation is always carried out over a contravariant index and its corresponding covariant index.

(iv) Free (nonrepeated) indices on one side of an equation must be balanced by the same indices on the other side of the equation.

(v) Contravariant components of a vector are denoted by superscripts; covariant components, by subscripts. The notation for the base vectors is just the opposite; subscripts denote contravariant base vectors, and superscripts denote covariant base vectors.

We have discussed the transformation between contravariant and covariant representations. We shall now consider the transformation of a contravariant or a covariant vector from one frame of reference to another. When two frames are in relative uniform motion, the transformation is determined by the Lorentz transformation laws. We can express the space–time coordinates of a physical event by either a contravariant or a covariant vector. The transformation from an unprimed frame to a primed frame is

$$x'^{\alpha} = P^{\alpha}_{\beta}x^{\beta} \tag{14a}$$

or

$$x'_\alpha = Q^\beta_\alpha x_\beta \tag{14b}$$

We can view P^α_β as a matrix, denoted by $\overline{\overline{P}}$, operating on column matrix x^β and giving column matrix x'^α. A similar view is applied to Q^β_α. In view of (7.1.1), the transformation matrices $\overline{\overline{P}}$ and $\overline{\overline{Q}}$ are as follows:

$$\overline{\overline{P}} = \begin{bmatrix} \gamma & -\gamma\beta_x & -\gamma\beta_y & -\gamma\beta_z \\ -\gamma\beta_x & 1+(\gamma-1)\beta_x^2/\beta^2 & (\gamma-1)\beta_x\beta_y/\beta^2 & (\gamma-1)\beta_x\beta_z/\beta^2 \\ -\gamma\beta_y & (\gamma-1)\beta_y\beta_x/\beta^2 & 1+(\gamma-1)\beta_y^2/\beta^2 & (\gamma-1)\beta_y\beta_z/\beta^2 \\ -\gamma\beta_z & (\gamma-1)\beta_z\beta_x/\beta^2 & (\gamma-1)\beta_z\beta_y/\beta^2 & 1+(\gamma-1)\beta_z^2/\beta^2 \end{bmatrix} \tag{15a}$$

$$\overline{\overline{Q}} = \begin{bmatrix} \gamma & \gamma\beta_x & \gamma\beta_y & \gamma\beta_z \\ \gamma\beta_x & 1+(\gamma-1)\beta_x^2/\beta^2 & (\gamma-1)\beta_x\beta_y/\beta^2 & (\gamma-1)\beta_x\beta_z/\beta^2 \\ \gamma\beta_y & (\gamma-1)\beta_y\beta_x/\beta^2 & 1+(\gamma-1)\beta_y^2/\beta^2 & (\gamma-1)\beta_y\beta_z/\beta^2 \\ \gamma\beta_z & (\gamma-1)\beta_z\beta_x/\beta^2 & (\gamma-1)\beta_z\beta_y/\beta^2 & 1+(\gamma-1)\beta_z^2/\beta^2 \end{bmatrix} \tag{15b}$$

A few properties of $\overline{\overline{P}}$ and $\overline{\overline{Q}}$ follow from (14). The magnitude squared of x^α is a Lorentz invariant. From

$$x'^\mu x'_\mu = P^\mu_\alpha Q^\beta_\mu x^\alpha x_\beta$$

we learn that

$$P^\mu_\alpha Q^\beta_\mu = \delta^\beta_\alpha$$

The summation is over μ. In matrix form, we have

$$\overline{\overline{P}}^t \cdot \overline{\overline{Q}} = \overline{\overline{I}}$$

where $\overline{\overline{I}}$ is the 4×4 unit matrix. Obviously, $\overline{\overline{P}}^t$ is the inverse of $\overline{\overline{Q}}$:

$$\overline{\overline{P}}^t = \overline{\overline{Q}}^{-1}$$

The inverse of $\overline{\overline{P}}$ is then the transpose of $\overline{\overline{Q}}$:

$$\overline{\overline{P}}^{-1} = \overline{\overline{Q}}^t$$

With these last two relations, the inverse transformation from primed to unprimed coordinates is as follows:

$$x^\alpha = Q^\alpha_\beta x'^\beta \tag{17a}$$

$$x_\alpha = P^\beta_\alpha x'_\beta \tag{17b}$$

When (17) is compared with (14), it can be verified that the two are equivalent by using (16).

Conventionally, a contravariant vector is defined as one that transforms with the transformation matrix $\overline{\overline{P}}$ as (14a); a covariant vector transforms with the matrix $\overline{\overline{Q}}$ as (14b). Extending the definition, an n th-rank contravariant tensor transforms from one Lorentz frame to another by using the transformation matrix $\overline{\overline{P}}$ n times. An n th-rank covariant tensor transforms from one Lorentz frame to another by using $\overline{\overline{Q}}$ n times. The scalar product of an n th-rank contravariant tensor with an n th-rank covariant tensor is Lorentz-invariant. For instance, the space–time derivatives $(\partial/\partial ct, \nabla)$ form a covariant four-vector because, according to (7.3.1), the transformation is like (14b) and (17b). If we denote the derivatives of a scalar function $\chi(x)$ by

$$\chi_{,\alpha} = (\partial\chi/\partial ct, \nabla\chi)$$

we find

$$\chi'_{,\alpha} = Q^{\beta}_{\alpha}\chi_{,\beta}$$

The charge current density, written as

$$J^{\alpha} = (c\rho, \overline{J}) \tag{18}$$

is seen to be a contravariant vector because it transforms as (14a) and (17a). The space–time derivative of J^{α}

$$J^{\alpha}{}_{,\alpha} = \frac{\partial\rho}{\partial t} + \nabla \cdot \overline{J} \tag{19}$$

becomes a scalar. The charge current conservation law states that

$$J^{\alpha}{}_{,\alpha} = 0 \tag{20}$$

From (11), we can show that $\eta_{\alpha\beta}$ is a second-rank covariant tensor and $\eta^{\alpha\beta}$ is a second-rank contravariant tensor. They are known as metric tensors.

The transformation matrices P^{α}_{β} and Q^{β}_{α} are pure Lorentz transformations. A pure Lorentz transformation satisfies two assumptions: (i) the coordinate axes of S and S' are parallel, and (ii) the origins of the two coordinate systems coincide at $t = 0$. A Lorentz transformation (LT) that satisfies (ii) but does not satisfy (i) is called a

homogeneous LT (HLT). A HLT is a combination of a pure LT plus a spatial rotation. Mathematically, the whole class of HLTs satisfies the postulates of a group and is called an HLT group. It is important to note that the group multiplication of two pure LTs results not in a pure LT, but in a pure LT plus a spatial rotation. The matrices P_β^α and Q_α^β in (14) can be used to represent the HLT group. Although as represented in (15) they appear to be symmetrical, symmetry is not a general property of all elements in the HLT group. For instance, the transformation matrix for a spatial rotation is not symmetrical. When assumption (ii) is violated, the LT is inhomogeneous. An inhomogeneous LT can be made homogeneous by re-choosing the time-space origins. The whole class of inhomogeneous LTs also forms a group called an inhomogeneous Lorentz group or simply the Poincaré group. The HLT group is a subgroup of the Poincaré group because the identity element is there. Any element in a Poincaré group can be joined to the identity element continuously by successive LTs. For completeness, we mention the LTs in which a spatial or space–time inversion is involved. This group of LTs is called an improper Lorentz group in which an element cannot continuously join to the identity. In our treatment, we confine ourselves to pure LTs in the HLT group.

b. Maxwell's Equations in Tensor Form

To write Maxwell's equations in compact tensor form, we define a field tensor $F_{\alpha\beta}$ and an excitation tensor $G_{\alpha\beta}$. Explicitly in terms of matrix representation, we have

$$F_{\mu\nu} = \begin{bmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -cB_z & cB_y \\ -E_y & cB_z & 0 & -cB_x \\ -E_z & -cB_y & cB_x & 0 \end{bmatrix} \tag{21}$$

$$G_{\mu\nu} = \begin{bmatrix} 0 & cD_x & cD_y & cD_z \\ -cD_x & 0 & -H_z & H_y \\ -cD_y & H_z & 0 & -H_x \\ -cD_z & -H_y & H_x & 0 \end{bmatrix} \tag{22}$$

Both tensors are skew-symmetric:

$$F_{\alpha\beta} = -F_{\beta\alpha}$$
$$G_{\alpha\beta} = -G_{\beta\alpha}$$

They are second-rank covariant tensors and transform as

$$F'_{\mu\nu} = Q^{\alpha}_{\mu} Q^{\beta}_{\nu} F_{\alpha\beta} \tag{23a}$$

$$G'_{\mu\nu} = Q^{\alpha}_{\mu} Q^{\beta}_{\nu} G_{\alpha\beta} \tag{23b}$$

Using matrix notation, we write

$$\overline{\overline{F}}' = \overline{\overline{Q}} \cdot \overline{\overline{F}} \cdot \overline{\overline{Q}}^{t}$$

$$\overline{\overline{G}}' = \overline{\overline{Q}} \cdot \overline{\overline{G}} \cdot \overline{\overline{Q}}^{t}$$

It is straightforward to show that the result of this transformation is identical to the Lorentz transformation which is obtained with three-dimensional notation.

The three-dimensional field vectors are related to the field tensor and the excitation tensor in the following manner:

$$E_i = F_{0i}$$

$$cB_i = -\epsilon_{ijk} F_{jk}$$

$$cD_i = G_{0i}$$

$$H_i = -\epsilon_{ijk} G_{jk}$$

The contravariant tensors corresponding to $F_{\alpha\beta}$ and $G_{\alpha\beta}$ can be obtained by using the matrix tensor $\eta^{\alpha\beta}$. For instance,

$$G^{\mu\nu} = \eta^{\mu\alpha} \eta^{\nu\beta} G_{\alpha\beta} = \begin{bmatrix} 0 & -cD_x & -cD_y & -cD_z \\ cD_x & 0 & -H_z & H_y \\ cD_y & H_z & 0 & -H_x \\ cD_z & -H_y & H_x & 0 \end{bmatrix} \tag{24}$$

This result can be obtained by the convention of raising and lowering indices. The contravariant components G^{0i} are negatives of their corresponding covariant components because an index 0 is raised.

In tensor notation, Maxwell's equations read as

$$F_{\alpha\beta,\gamma} + F_{\beta\gamma,\alpha} + F_{\gamma\alpha,\beta} = 0 \tag{25a}$$

$$G^{\alpha\mu}_{,\alpha} = J^{\mu} \tag{25b}$$

We now demonstrate that (25b) is equivalent to the full set of Maxwell's equations. If none of the three indices α, β, γ in (25a) is zero, we find Gauss' law of magnetism:

$$\nabla \cdot \overline{B} = 0$$

If one of α, β, γ is zero, $(25a)$ gives Faraday's law:

$$\nabla \times \overline{H} + \frac{\partial \overline{B}}{\partial t} = 0$$

If $\mu = 0$ in $(25b)$, we obtain Gauss' law of electricity:

$$\nabla \cdot \overline{D} = \rho$$

For $\mu \neq 0$, $(25b)$ gives

$$\nabla \times \overline{E} - \frac{\partial \overline{D}}{\partial t} = \overline{J}$$

which is Ampere's law with the Maxwell's displacement current term $\partial \overline{D}/\partial t$.

The conventional exercise of expressing field vectors in terms of vector and scalar potentials is observed from $(25a)$. It is quite easy to show that $(25a)$ is satisfied if

$$F_{\alpha\beta} = A_{\alpha,\beta} - A_{\beta,\alpha} \tag{26}$$

where A_α is a covariant four-vector, its zeroth contravariant component is the scalar potential ϕ, and its space components are the vector potential $\overline{A}$ times c:

$$A_\alpha = \begin{bmatrix} -\phi \\ c\overline{A} \end{bmatrix}, \qquad A^\alpha = \begin{bmatrix} \phi \\ c\overline{A} \end{bmatrix} \tag{27}$$

Writing (26) in three-dimensional notation, we have the familiar expressions

$$\overline{E} = -\frac{\partial \overline{A}}{\partial t} - \nabla\phi$$
$$\overline{B} = \nabla \times \overline{A}$$

If we make a gauge transformation from A_α to A'_α so that

$$A_\alpha = A'_\alpha + \psi_{,\alpha} \tag{28}$$

where ψ is any scalar function of time–space by introducing (28) in $(25a)$, we have

$$F_{\alpha\beta} = A_{\alpha,\beta} - A_{\beta,\alpha} = A'_{\alpha,\beta} - A'_{\beta,\alpha} \tag{29}$$

This shows that both A_α and A'_α give rise to the same field tensor. This arbitrariness is fixed by the gauge condition. The Lorentz gauge is

$$A^{\mu}_{,\mu} = 0 \tag{30}$$

which takes the same form as the continuity equation for charge current densities. Unlike the Coulomb gauge, the Lorentz gauge is relativistically covariant.

c. *Constitutive Relations in Tensor Form*

The constitutive relations in tensor notation provide a relation for the excitation tensor $G^{\alpha\beta}$ and the field tensor $F_{\alpha\beta}$. We write

$$G^{\alpha\beta} = \frac{1}{2} C^{\alpha\beta\rho\sigma} F_{\rho\sigma} \tag{31}$$

We call the fourth-rank tensor $C^{\alpha\beta\rho\sigma}$ the constitutive tensor. Because of the skew-symmetric properties of $F_{\rho\sigma}$ and $G^{\alpha\beta}$, we see that

$$C^{\alpha\beta\rho\sigma} = -C^{\beta\alpha\rho\sigma} = -C^{\alpha\beta\sigma\rho} = C^{\beta\alpha\sigma\rho} \tag{32}$$

The constitutive tensor is skew-symmetric with respect to the first pair, as well as the second pair, of indices. In general, a fourth-rank tensor in a four-dimensional space possesses 256 elements. Because of this skew symmetry, the first pair of indices has six independent elements and so does the second pair, giving rise to a total of 36 independent elements. Thus the 6×6 constitutive matrix $\overline{\overline{C}}$ is a faithful representation of the constitutive tensor.

We shall establish relations between the tensor elements of $C^{\alpha\beta\rho\sigma}$ and the matrix elements of $\overline{\overline{C}}$ in (31). In view of (32), we have

$$C^{0i0j} = p_{ij} \tag{33a}$$

$$C^{ijkl} = \epsilon_{ijm}\epsilon_{kln} q_{mn} \tag{33b}$$

$$C^{0kij} = -\epsilon_{ijm} l_{km} \tag{33c}$$

$$C^{ij0k} = -\epsilon_{ijn} m_{kn} \tag{33d}$$

In Chapter II we demonstrated that symmetric conditions exist for lossless media where $p_{ij} = p^*_{ji}$, $q_{mn} = q^*_{nm}$, and $l_{kn} = -m^*_{nk}$ which

limit the independent constitutive parameters to 21. In terms of the constitutive tensor, the lossless conditions correspond to

$$C_{\rho\sigma}^{\alpha\beta} = (C_{\rho\sigma}^{\alpha\beta})^* \tag{34}$$

Note that, if we were to express this in terms of four contravariant components, we would have $C^{0kij} = -\epsilon_{imj}l_{km} = \epsilon_{ijm}m_{km}^* = -(C^{ij0k})^*$ while $C^{\alpha i\beta j} = (C^{\beta j\alpha i})^*$. We obtain the final result using two contravariant and two covariant indices.

At this point it is appropriate to mention the work of Tischer and Hess [1959] on a covariant description of a conducting medium. Ohm's law relates the conduction current to the electric field by conductivity, which can be isotropic as well as anisotropic. For moving anisotropic conducting media, Tischer and Hess introduced a new three-dimensional vector, together with the conducting current, to form a four-dimensional skew-symmetric tensor just like $F_{\mu\nu}$ and $G_{\mu\nu}$. They thus obtained a covariant description of what we may call bianisotropic conducting media. The implications of the new vector have not been explored.

Other covariant descriptions of moving isotropic media exist, such as

$$(\delta_\rho^\alpha + u^\alpha u_\rho)J_c^\rho = \sigma u_\beta F^{\alpha\beta}$$

for Ohm's law and

$$G_{\lambda\mu} = F_{\lambda\mu} + (n^2 - 1)(F_{\mu\sigma}u^\sigma u_\lambda - F_{\lambda\sigma}u^\sigma u_\mu)$$

for isotropic nonconducting media. The velocity four-vector $u^\alpha = (1, \overline{\beta})$. Thus the manifestly covariant descriptions explicitly display the velocity dependence. When reference is made to the rest frame of the medium, $u^\alpha = (1,0,0,0)$ and the two equations yield $\overline{J} = \sigma\overline{E}$, $\overline{D} = \epsilon\overline{E}$, and $\overline{H} = \overline{B}/\mu$.

7.11 Hamilton's Principle and Noether's Theorem

a. Action Integral

Starting from a postulated Lagrangian density, the variational principle provides an elegant and systematic way of deriving the equations of motion and the conservation laws of a physical system. In the

case of macroscopic electromagnetic fields the Lagrangian density is postulated as

$$\begin{aligned} L[x_\alpha, A_\alpha(x_\mu), A_{\alpha,\beta}(x_\mu)] &= -\frac{1}{4} F^{\alpha\beta} G_{\alpha\beta} + J^\alpha A_\alpha \\ &= -\frac{1}{8} C^{\alpha\beta\rho\sigma} (A_{\alpha,\beta} - A_{\beta,\alpha})(A_{\rho,\sigma} - A_{\sigma,\rho}) + J^\alpha A_\alpha \end{aligned} \tag{1}$$

The Lagrangian density $L(x, A_\alpha, A_{\alpha,\beta})$ is a function of the space–time coordinates x_α, the potential functions A_α, and the space–time derivatives of the potential functions $A_{\alpha,\beta}$. The potential functions A_α are also called state functions. The charge current four-vector J_α is an externally given state function.

The variational principle applies to an action integral I, defined by integration of the Lagrangian density over a four-dimensional space R:

$$I = \int_R d^4x \; L[x_\alpha, A_\alpha(x_\mu), A_{\alpha,\beta}(x_\mu)] \tag{2}$$

The variation of the action integral is caused by either a variation of the state functions A_α or a variation of the domain of integration R which induces variations on the space–time-dependent state functions A_α, and on the externally given J_α because both are space–time dependent.

b. Hamilton's Principle and Maxwell's Equations

In Hamilton's principle, the domain of integration R is not varied. The state function A_α inside the domain R is varied by an arbitrary and infinitesimally small amount δA_α:

$$A'_\alpha(x_\mu) = A_\alpha(x_\mu) + \delta A_\alpha(x_\mu) \tag{3}$$

where A'_α are new state functions. The state functions on the boundary of R are not varied, where $\delta A_\alpha = 0$. The principle requires that the action integral be stationary under such variations:

$$\delta I = \int_R d^4x \, [L'(x_\alpha, A'_\alpha, A'_{\alpha,\beta}) - L(x_\alpha, A_\alpha, A_{\alpha,\beta})] = 0 \tag{4}$$

The new Lagrangian density $L'(x_\alpha, A'_\alpha, A'_{\alpha,\beta})$ is related to the old

Lagrangian density L as follows:

$$\begin{aligned} L'(x_\alpha, A'_\alpha, A'_{\alpha,\beta}) &= L(x_\alpha, A_\alpha, A_{\alpha,\beta}) + \frac{\partial L}{\partial A_\alpha}\delta A_\alpha + \frac{\partial L}{\partial A_{\alpha,\beta}}\delta A_{\alpha,\beta} \\ &= L(x_\alpha, A_\alpha, A_{\alpha,\beta}) + \frac{\partial L}{\partial A_\alpha}\delta A_\alpha \\ &\quad - \frac{d}{dx^\beta}\left(\frac{\partial L}{\partial A_{\alpha,\beta}}\right)\delta A_\alpha + \frac{d}{dx^\beta}\left(\frac{\partial L}{\partial A_{\alpha,\beta}}\delta A_\alpha\right) \end{aligned}$$

Substituting in (4) yields

$$\int_R d^4x \left\{\left[\frac{\partial L}{\partial A_\alpha} - \frac{d}{dx^\beta}\left(\frac{\partial L}{\partial A_{\alpha,\beta}}\right)\right]\delta A_\alpha + \frac{d}{dx^\beta}\left(\frac{\partial L}{\partial A_{\alpha,\beta}}\delta A_\alpha\right)\right\} = 0 \tag{5}$$

The last term vanishes after integration because of the condition that δA_α vanishes on the boundary of R. Since δI vanishes for all variations of the state function δA_α, we have

$$\frac{\partial L}{\partial A_\alpha} - \frac{d}{dx^\beta}\left(\frac{\partial L}{\partial A_{\alpha,\beta}}\right) = 0 \tag{6}$$

which is known as the *Euler-Lagrangian equation.*

Maxwell's equation

$$F_{\alpha\beta,\gamma} + F_{\beta\gamma,\alpha} + F_{\gamma\alpha,\beta} = 0 \tag{7}$$

is a direct consequence of the definition

$$F_{\alpha\beta} = A_{\alpha,\beta} - A_{\beta,\alpha} \tag{8}$$

The other Maxwell's equation,

$$G^{\beta\alpha}_{\ ,\beta} = J^\alpha \tag{9}$$

can be derived from the Euler-Lagrangian equation by using the equations $\partial L/\partial A_\alpha = J^\alpha$ and $\partial L/\partial A_{\alpha,\beta} = -G^{\alpha\beta}$.

c. Noether's Theorem and Energy Momentum Tensors

In Noether's theorem the domain of integration R is rendered an infinitesimal transformation, which induces variations on both state

functions A_α and J_α. Suppose that the domain R is mapped onto a new domain R' such that

$$x'_\alpha = x_\alpha + \delta x_\alpha \tag{10}$$

This mapping transplants a state function from x_α to x'_α :

$$A'_\alpha(x') = A_\alpha(x) + \delta A_\alpha \tag{11}$$

Note that δA_α gives the difference between the new state function $A'_\alpha(x')$ at the new location x'_α and the old state function $A_\alpha(x)$ at the old location x_α before it is transplanted. Usually the state functions are not explicit functions of space–time coordinates. Consider an infinitesimal translation of the domain R :

$$x'_\alpha = x_\alpha + \epsilon_\alpha \tag{12}$$

From the active viewpoint of a coordinate transformation, all state functions are transplanted by an infinitesimal amount. The new and the old state functions are equal in magnitude, and there is no change in the orientation. From the passive viewpoint of coordinate transformation, the coordinate axes are translated by an infinitesimal amount and the state functions are left unchanged. We have

$$\delta A_\alpha = 0 \tag{13}$$

Next, consider an infinitesimal rotation of the coordinate axes:

$$x'_\alpha = x_\alpha - \omega_\alpha^\beta x_\beta \tag{14}$$

The induced variations on the state functions are

$$\delta A_\alpha = \omega_\alpha^\beta A_\beta \tag{15}$$

From the active viewpoint, the state functions are rotated by an infinitesimal amount. From the passive viewpoint, the amount is equal to the change of the component projections on the new and the old coordinate axes.

The variation denoted by δA_α is rather awkward because it compares two state functions at different locations. We define instead

$$\overline{\delta} A_\alpha = A'_\alpha(x') - A_\alpha(x') = A'_\alpha(x) - A_\alpha(x) \tag{16}$$

to denote the difference between the new and the old state functions at the same location. The second equality is valid up to the first order, as shown here:

$$\begin{aligned}\bar{\delta}A_\alpha &= A'_\alpha(x') - A_\alpha(x') \\ &= A'_\alpha(x) - A_\alpha(x) + (A'_{\alpha,\beta} - A_{\alpha,\beta})\delta x^\beta + \cdots \\ &\approx A'_\alpha(x) - A_\alpha(x)\end{aligned} \tag{17}$$

As a matter of fact, all first-order quantities with infinitesimal space–time separation can be shown to be equal. The relation between δA_α and $\bar{\delta}A_\alpha$ is easily established:

$$\begin{aligned}\bar{\delta}A_\alpha &= A'_\alpha(x') - A_\alpha(x') \\ &= A'_\alpha(x') - A_\alpha(x) - [A_\alpha(x') - A_\alpha(x)] \\ &= \delta A_\alpha - A_{\alpha,\beta}\delta x^\beta\end{aligned} \tag{18}$$

Note that, when the domain variation is a translation, $\delta A_\alpha = 0$ because the new state function is the same as the old one. But $\bar{\delta}A_\alpha \neq 0$ because the new state function at x'_α is transplanted from a neighboring location and is certainly different from the old state function at x'_α, which has been transplanted to another location.

Under the infinitesimal domain variations, variations are induced on both the state functions A_α, $A_{\alpha,\beta}$ and the externally given J_α. The new Lagrangian density is related to the old one by

$$\begin{aligned}L'(x'_\alpha, A'_\alpha, A'_{\alpha,\beta}) &= L + L_{,\alpha}\delta x^\alpha + \frac{\partial L}{\partial A_\alpha}\delta A_\alpha + \frac{\partial L}{\partial A_{\alpha,\beta}}\delta A_{\alpha,\beta} + \frac{\partial L}{\partial J_\alpha}\delta J_\alpha \\ &= L + \frac{dL}{dx^\rho}\delta x^\rho + \frac{\partial L}{\partial A_\alpha}\bar{\delta}A_\alpha + \frac{\partial L}{\partial A_{\alpha,\beta}}\bar{\delta}A_{\alpha,\beta} + \frac{\partial L}{\partial J_\alpha}\bar{\delta}J_\alpha\end{aligned}$$

where

$$\frac{dL}{dx^\rho} = L_{,\rho} + \frac{\partial L}{\partial A_\alpha}A_{\alpha,\rho} + \frac{\partial L}{\partial A_{\alpha,\beta}}A_{\alpha,\beta\rho} + \frac{\partial L}{\partial J_\alpha}J_{\alpha,\rho}$$

The new domain of integration R' is related to the old domain of integration by the Jacobian, which is

$$\det\left|\frac{\partial x'^\alpha}{\partial x^\beta}\right| = 1 + (\delta x^\alpha)_{,\alpha} \tag{19}$$

The variation of the action integral under this domain variation becomes, to the first order,

$$\begin{aligned}\delta I &= \int_{R'} d^4x'\, L'(x'_\alpha, A'_\alpha, A'_{\alpha\beta}) - \int_R d^4x\, L(x_\alpha, A_\alpha, A_{\alpha,\beta}) \\ &= \int_R d^4x \left\{ [1 + (\delta x^\alpha)_{,\alpha}] \left(L + \frac{dL}{dx^\rho}\delta x^\rho + \frac{\partial L}{\partial A_\alpha}\bar{\delta}A_\alpha + \frac{\partial L}{\partial J_\alpha}\bar{\delta}J_\alpha \right.\right. \\ &\qquad \left.\left. + \frac{\partial L}{\partial A_{\alpha,\beta}}\bar{\delta}A_{\alpha,\beta} \right) - L \right\} \\ &= \int_R d^4x \left\{ \left[\frac{\partial L}{\partial A_\alpha} - \frac{d}{dx^\beta}\left(\frac{\partial L}{\partial A_{\alpha,\beta}} \right) \right] \right. \\ &\qquad \left. + \frac{d}{dx^\rho}\left[L\delta x^\rho + \frac{\partial L}{\partial A_{\alpha,\rho}}\bar{\delta}A_\alpha \right] \bar{\delta}A_\alpha + \frac{\partial L}{\partial J_\alpha}\bar{\delta}J_\alpha \right\}\end{aligned}$$

Note that, although this variation of the action integral is derived from induced variations caused by infinitesimal domain changes, the result is fairly general and includes Hamilton's principle as a special case. Let us keep the domain unchanged; we then have $\delta x^\rho = 0$, $\bar{\delta}A_\alpha = \delta A_\alpha$. The externally given J_α is also not varied: $\delta J_\alpha = 0$. This result is seen to reduce to (5).

Noether's theorem requires that the action integral be stationary, and that the Euler-Lagrangian equations be satisfied for the state functions A_α under the domain variations. As a result, we obtain

$$\frac{d}{dx^\rho}\left[L\delta x^\rho + \frac{\partial L}{\partial A_{\alpha,\rho}}\bar{\delta}A_\alpha \right] + \frac{\partial L}{\partial J_\alpha}\bar{\delta}J_\alpha = 0 \tag{20}$$

This equation gives all of the conservation laws. We shall first consider the case of translation, which yields energy momentum tensors for macroscopic electromagnetic fields. The case of four-dimensional rotation is then considered, and the angular momentum conservation laws are derived. Under an infinitesimal translation, we have

$$\bar{\delta}A_\alpha = -A_{\alpha,\beta}\epsilon^\beta \tag{21a}$$

$$\bar{\delta}J_\alpha = -J_{\alpha,\beta}\epsilon^\beta \tag{21b}$$

Equation (20) gives

$$\epsilon^\beta \left[\frac{d}{dx^\rho}\left(-\frac{1}{4}F^{\mu\nu}G_{\mu\nu}\delta^\rho_\beta + J^\alpha A_\alpha \delta^\rho_\beta + G^{\alpha\rho}A_{\alpha,\beta} \right) - A^\alpha J_{\alpha,\beta} \right] = 0$$

This equation can be expressed in terms of field variables alone. After some manipulation and by use of the fact that $G^{\alpha\rho}A_{\beta,\alpha\rho} = 0$, we eliminate the potentials and obtain

$$T^{\alpha\beta}_{,\alpha} = -f^{\beta} \tag{22}$$

where

$$T^{\alpha\beta} = -\frac{1}{4}\eta^{\alpha\beta}F_{\rho\sigma}G^{\rho\sigma} + \eta^{\mu\beta}G^{\rho\alpha}F_{\rho\mu} \tag{23}$$

is the four-dimensional energy momentum tensor, and

$$f^{\beta} = J_{\alpha}F^{\alpha\beta} \tag{24}$$

In three-dimensional vector notation, we find that

$$f^{\beta} = \begin{bmatrix} \overline{J}\cdot\overline{E} \\ c\rho\overline{E} + \overline{J}\times c\overline{B} \end{bmatrix} \tag{25}$$

$$T^{\alpha\beta} = \begin{bmatrix} cW & c^2\overline{G} \\ \overline{S} & c\overline{\overline{T}} \end{bmatrix} \tag{26}$$

where

Electromagnetic energy	$W = \frac{1}{2}(\overline{D}\cdot\overline{E} + \overline{B}\cdot\overline{H})$
Energy flow density	$\overline{S} = \overline{E}\times\overline{H}$
Momentum density	$\overline{G} = \overline{D}\times\overline{B}$
Maxwell stress tensor	$\overline{\overline{T}} = \frac{1}{2}(\overline{D}\cdot\overline{E} + \overline{B}\cdot\overline{H})\overline{\overline{I}} - \overline{D}\,\overline{E} - \overline{B}\,\overline{H}$

The conservation law (22), written in vector notation, takes the form

$$\nabla\cdot\overline{S} + \frac{\partial W}{\partial t} = -\overline{J}\cdot\overline{E} \tag{27a}$$

$$\nabla\cdot\overline{\overline{T}} + \frac{\partial\overline{G}}{\partial t} = -(\rho\overline{E} + \overline{J}\times\overline{B}) \tag{27b}$$

The first equation is Poynting's theorem. The second equation was derived also in 1.3. Under an infinitesimal rotation, we have

$$\overline{\delta}A_{\alpha} = \delta A_{\alpha} - A_{\alpha,\beta}\delta x^{\beta} = \omega^{\beta}_{\alpha}A_{\beta} - A_{\alpha,\beta}\omega^{\beta}_{\rho}x^{\rho} \tag{28a}$$

$$\overline{\delta}J_{\alpha} = \omega^{\beta}_{\alpha}J_{\beta} - J_{\alpha,\beta}\omega^{\beta}_{\rho}x^{\rho} \tag{28b}$$

After some manipulations, (20) gives

$$\omega_{\alpha\beta}\left[\frac{d}{dx^\rho}(T^{\rho\alpha}x^\beta + G^{\alpha\rho}A^\beta + G^{\beta\rho}A^\alpha) + J_\rho F^{\rho\alpha}x^\beta + \eta^{\alpha\beta}J^\rho A_\rho\right] = 0 \tag{29}$$

Interchanging α and β, we have

$$\omega_{\alpha\beta}\left[\frac{d}{dx^\rho}(T^{\rho\alpha}x^\beta + G^{\alpha\rho}A^\alpha + G^{\beta\rho}A^\beta) + J_\rho F^{\rho\alpha}x^\alpha + \eta^{\beta\alpha}J^\rho A_\rho\right] = 0 \tag{30}$$

We add the two equations above and note that $\omega_{\beta\alpha} = -\omega_{\alpha\beta}$, obtaining

$$\frac{d}{dx^\rho}M^{\rho\alpha\beta} = x^\alpha J_\rho F^{\rho\beta} - x^\beta J_\rho F^{\rho\alpha} \tag{31}$$

where

$$M^{\rho\alpha\beta} = T^{\rho\alpha}x^\beta - T^{\rho\beta}x^\alpha \tag{32}$$

is the four-dimensional angular momentum tensor for electromagnetic fields.

PROBLEMS

Problem P7.1

To consider transformation of time intervals, let a clock be in S', which moves along the $\hat{z}$ direction of S. The time interval of the clock as read by S' is $\Delta t' = t'_2 - t'_1$ and is called the *proper time interval.* The corresponding time interval of the clock as read by S is $t = t_2 = t_1$ and is called the *coordinate time interval.* Show that

$$c\Delta t' = \gamma(c\Delta t - \beta\Delta z)$$
$$\Delta z' = \gamma(\Delta z - \beta c\Delta t)$$

The clock is stationary in S', and hence $\Delta z' = 0$. Show that $\Delta t = \gamma\Delta t'$. Observe that the coordinate time interval is always larger than the proper time interval. This phenomenon is known as time dilation.

Problem P7.2

Assume that S' moves with velocity v_1 relative to S and S'' moves with velocity v_2 relative to S, both along the z axis of S.

Write the LT between S' and S, and the LT between S'' and S. Eliminate the space–time coordinates z and t of S from the two LTs and show that the resulting LT implies a relative velocity u between S'' and S', where

$$u = \frac{(v_2 - v_1)}{(1 - v_1 v_2/c^2)}$$

This is an additive law for two velocities along the same direction. Generalize this procedure and deduce an additive law for two velocities in different directions.

Problem P7.3

In the early stages of special relativity, the sudden disappearance of an absolute time scalar led to the well-known "twin paradox." The paradox as stated was that one of a pair of twins left home, traveled at a uniform (high) speed in some direction for a certain period of time, and then returned home to find himself younger than his brother. By the symmetry argument that motion is relative, it was argued that neither twin should have grown older than the other, and the validity of special relativity was challenged. In the following discussion we show that both of the twins agree that one is older than the other and the problem is not symmetric.

Let both A and B be at the origin in frame S; B starts to move at $t = 0$ with speed v in the positive $\hat{z}$ direction of S. As A reads time t, B moves back with speed v. Consider the following events:

Event 1: Twin B is at $z = vt$ when A reads time t. In frame S, this event is described by (ct, vt) where ct is the time coordinate of the event with the dimension length, and vt is the space coordinate of the event.

Event 2: As A reads time $2t$, both A and B are at $z = 0$. In frame S, this event is described by $(2ct, 0)$.

Consider two other frames of reference, S' and S''. Frame S' moves with velocity v in the positive direction of the z axis. Show that the space–time coordinates for the two events in frames S' and S'' are as listed in Table P7.3.

(a) Using the space–time coordinates of the two events, show that, by time dilation, twin A agrees that the total proper time interval for twin B is $2t/\gamma$, while his own coordinate time interval is $2t$.

Event	Observer		
	S	S'	S''
1	$ct(1,\beta)$	$ct(1/\gamma,0)$	$ct[\gamma(1+\beta^2),2\gamma\beta]$
2	$ct(2,0)$	$ct(2\gamma,-2\gamma\beta)$	$ct(2\gamma,2\gamma\beta)$

Table P7.3 Time–space coordinates of the two events in the three frames of reference. The first part in parentheses denotes time coordinates multiplied by c, and the second part denotes space coordinates.

(b) During the initial period before turning around, B is in S', and the elapsed time according to B is t/γ. Show that, according to B, the elapsed time during the final period after turning around is also t/γ. Thus twin B agrees with twin A that his time space is $2t/\gamma$, while that of A is $2t$.

(c) Show that observers moving uniformly relative to A, especially those in frame S' and S'', conclude that the proper elapsed time of B is less than that of A by a factor of $1/\gamma$.

(d) Suppose that twin B started his journey right after his birth and travelled with a speed $v = 0.8c$. If he comes back at 30 years of age, how old is his twin A?

(e) The problem is inherently asymmetrical; one twin has to turn, and it is this twin that experiences less proper time, $2t/\gamma$. If B does not turn, then, as A reads proper time $2t$, the coordinate time reading for B at two different locations, $z = 0$ and $z = 2vt$, is $2\gamma t$ because of the dilation of time. After turning around and meeting A again at $z = 0$, the proper time reading of B has been shown to be $2t/\gamma$. The effect of turning around causes a time difference of $2t(\gamma - 1/\gamma) = 2\gamma\beta^2 t$. Show that this "lost time" is equal to the time coordinate difference between S' and S'' for Event 1 in Table P7.3.

(f) It is interesting to imagine how B experiences the period of losing time during his turning around. Consider a third event, occurring when A reads time t at $z = 0$ in S. Find the time–space coordinates for Event 3 in S' and S''. Show that, according to S, Events 1 and 3 are simultaneous; according to S', Event 1 is earlier than Event 3; and according to S'', Event 3 is earlier than Event 1. At the turning time, twin B changes his frame from S'

to S''. Show that B loses track of anything that happens at $z = 0$ during a time period $2\gamma\beta^2 t$.

Problem P7.4

The star Alpha Centauri is 4.3 light-years from Earth. Observer B leaves Earth in a rocket ship that travels toward this star at acceleration g. Halfway (2.15 light-years from Earth) from α Cen, B turns off the forward acceleration and accelerates backward toward Earth at g, so that the rocket arrives at α Cen with zero speed and turns back. On the return trip, at the halfway point, B again changes the direction of acceleration. Observer B arrives at Earth with zero speed. Show that B takes 7 years to complete the round trip, while the elapsed time on Earth is 12 years.

Problem P7.5

Show that a rigid rod moving along its longitudinal direction with velocity v appears to be shortened by a factor $1/\gamma$. This phenomenon is known as the *Lorentz contraction*. Let the rod be at rest in S'. Its two end points are $z' = 0$ and $z' = z'_1$. In the laboratory frame S, the rod is moving along z with velocity v. Its length is measured by recording the positions of its end points simultaneously, $t_2 = t_1$. Show that, by LT, $\Delta z = z'_1/\gamma$, where $\Delta z = z_2 - z_1$ is the rod length as measured in S. Note that from the point of view of observer S', who moves with the rod, the laboratory observer is not measuring the two end points simultaneously. Consider the end points from two events.

Show that, according to S', S measures z_2 at $ct' = -\gamma\beta z_2$ and z_1 at $t' = 0$. For a rod moving from left to right, if we measure the position of the right end point first and the left end point next, we naturally obtain a result smaller than the actual length of the rod. Thus S' expects that S will claim a shorter length.

Problem P7.6

A rigid square board with two sides parallel to the z axis and two sides perpendicular to the z axis is moving uniformly along the z axis with respect to an observer S at the origin. Do you expect S to observe a rectangular board because of the effect of the Lorentz contraction? Show that, in fact, as viewed by observer S, the board is still square but is turned by an angle resulting from the relative motion.

Event	Observer	
	S	S'
1	$[t_0, \overline{P}(t_0)]$	$[t'_0, \overline{P}'(t'_0)]$
2	$[t_0, \overline{Q}(t_0)]$	$[t', \overline{Q}'(t')]$

Table P7.7 Time–space coordinates of two end points.

Problem P7.7

Consider the transformation from observer S' to observer S of the length of a rod that may or may not be rigid. From the point of view of S, the length of the rod at time t_0 is defined as the difference between space positions of the two end points as measured simultaneously at t_0 in his frame of reference

$$\overline{l}(t_0) = \overline{Q}(t_0) - \overline{P}(t_0)$$

where $\overline{Q}(t_0)$ and $\overline{P}(t_0)$ denote position readings of the two end points. But a simultaneous measurement at two space points in S is not simultaneous in S'. Consider the general case when the length of the rod is time-variant, and let the measurement performed be represented by two events. Their space–time coordinates are listed in Table P7.7.

By using the space–time Lorentz transformation formulas, show that

$$t' - t'_0 = -\overline{\beta} \cdot \overline{X}$$
$$\overline{l} = \overline{\alpha}^{-1} \cdot \overline{X}$$
$$\overline{X} = \overline{Q}'(t') - \overline{P}'(t'_0)$$

Note that, when the rod is not rigid and stationary in S', $\overline{X}$ is not the result of a length measurement performed in S' because the two end points are not measured simultaneously and hence do not conform to the definition of length.

(a) Show that, when the rod is rigid and stationary in S', the rod as viewed from S has length

$$\overline{l} = \overline{\overline{\alpha}}^{-1} \cdot \overline{l}' = \frac{1}{\gamma} \overline{l}_{\|} + \overline{l}'_{\perp}$$

where subscripts $\|$ and $\perp$ denote components of a vector parallel and perpendicular, respectively, to the velocity. Thus the parallel component is unchanged. The rod appears to be shortened and rotated by an angle.

(b) When the rod is rigid and is moving uniformly with velocity $\vec{v}'$ in S', the length of the rod in S' at t_0' is

$$\vec{l}'(t_0') = \vec{Q}'(t_0') - \vec{P}(t_0')$$

We can write $\vec{X}$ in terms of l' and $\vec{v}'$:

$$\vec{X} = \vec{v}'(t' - t_0') + \vec{l}'$$

Show that, in terms of $\vec{l}'$,

$$\vec{l} = \overline{\overline{\alpha}}^{-1} \cdot \left[\vec{l}' - \vec{\beta}' \frac{\vec{\beta} \cdot \vec{l}'}{1 + \vec{\beta} \cdot \vec{\beta}} \right]$$

where $\vec{\beta}' = \vec{v}'/c$. Thus, together with the Lorentz contraction, there is another change of length in the rigid rod because of its motion in S'.

(c) The rod is short and is time-variant. Since it is short, we can expand $t' - t_0$ to

$$t' - t_0' = \sum_{l=1}^{n} A_l \delta^l$$

where $\delta = \vec{\beta} \cdot \vec{l}'$ denotes a small quantity. Expand $\vec{X}$ to the nth order in δ, and show that

$$\begin{aligned} \vec{X} &= \vec{l}' + \sum_{k=1}^{\infty} \frac{\vec{Q}'^{(k)}(t_0')}{k!} (t' - t_0')^k \\ &= \vec{l}' + \sum_{k=1}^{N} \frac{\vec{Q}'^{(k)}}{k!} \left[\sum_{l=1}^{n} A_l \delta^l \right]^k \end{aligned}$$

where $\vec{Q}'^{(k)}(t_0)$ is the kth derivative of $\vec{Q}'(t')$ relative to t' at $t' = t_0'$. Comparing coefficients for δ^k, $k = 1, 2, \ldots, n$, determine

the coefficients A_k in the following way:

$$k=1: \quad (1+\overline{\beta}\cdot\overline{Q}')A_1=-1$$
$$k=2: \quad (1+\overline{\beta}\cdot\overline{D}')A_2=-\frac{1}{2}(\overline{\beta}\cdot\overline{Q}'^{(2)})A_1^2$$
$$k=3: \quad (1+\overline{\beta}\cdot\overline{D}')A_3=1(\overline{\beta}\cdot\overline{Q}'^{(2)})A_1A_2-\frac{1}{3!}(\overline{\beta}\cdot\overline{Q}'^{(3)})A_1^3$$

Prove that, to the second order in δ, $\bar{l}$ reads as

$$\bar{l}=\overline{\alpha}^{-1}\cdot\left(\overline{\overline{I}}+A_1\overline{D}'^{-}\right)\left[\overline{\overline{I}}+\frac{1}{2}A_1^2\delta\left(1-A_1^2\delta\overline{\beta}\cdot\overline{Q}'^{(2)}\right)\overline{Q}'^{(2)}\overline{\beta}\right]\cdot\bar{l}^{-1}$$

and, to the third order in δ,

$$\bar{l}=\overline{\alpha}^{-1}\cdot\left(\overline{\overline{I}}+A_1\overline{D}'^{-}\overline{\beta}\right)\cdot\left[\overline{\overline{I}}+\frac{1}{2}A_1^2\delta\left(1-A_1^2\delta\overline{\beta}\cdot\overline{Q}'^{(2)}+\frac{1}{3}A_1^3\delta^2\overline{\beta}\cdot\overline{Q}'\right)\overline{Q}'^{(2)}\overline{\beta}\right]\cdot\bar{l}^{-1}$$

Problem P7.8

Derive the transformation formula for $\overline{E}$ and $c\overline{B}$ as shown in (7.3.12). Find the corresponding FOLT and GT formulas. What are the inverse transformations of these formulas?

Problem P7.9

A positive charge moves along the z axis in frame S. Using transformation law (7.3.12), find the $\overline{B}$ field due to the moving charge. Does your result agree with Ampere's law?

Problem P7.10

Show that a moving current loop generates an electric dipole moment by using transformation law (7.3.12). Indicate the polarization of the dipole. For simplicity, consider a square loop.

Problem P7.11

In a certain reference frame, a static uniform electric field E_0 is parallel to the z axis and a static uniform magnetic field $cB_0 = 2E_0$ forms a 30° angle with respect to $\hat{z}$. Determine the relative velocity of a reference frame in which the electric and magnetic fields are parallel.

Problem P7.12

An observer S observes a uniform electric field in the $\hat{x}$ direction, $\overline{E} = \hat{x}E_0$, and a uniform magnetic field in the $\hat{y}$ direction, $\overline{B} = \hat{y}B_0$. Let $E_0 > cB_0$. Find an observer S' moving relative to S with velocity v in the $\hat{z}$ direction, so that he observes only an electric field. Determine the electric field strength and the velocity v. Can you find an observer moving with a velocity less than c who observes only a magnetic field?

Problem P7.13

What are the constitutive relations for a moving biisotropic medium and for a biaxial medium moving along one of its principal axes?

Problem P7.14

Consider a plane wave normally incident upon a dielectric medium moving toward the wave. Let the boundary be moving in the $\hat{z}$ direction with velocity v, and the electric field be linearly polarized in the $\hat{x}$ direction. Write

$$\overline{E}_i = \hat{x}E_0 e^{ikz - i\omega t}, \qquad k = -\frac{\omega}{c}$$

(a) Show that the fields for the reflected and transmitted waves are

$$\overline{E}_r = \hat{x}RE_0 e^{ik_r z - i\omega_r t}, \qquad k_r = \frac{\omega_r}{c}$$

$$\overline{E}_t = \hat{x}TE_0 e^{ik_t z - i\omega_t t}, \qquad k_t = \frac{n-\beta}{n\beta - 1}\frac{\omega_t}{c}$$

$$c\overline{B}_r = \hat{y}\frac{ck_r}{\omega_r}E_r = \hat{y}E_r$$

$$c\overline{B}_t = \hat{y}\frac{ck_t}{\omega_t}E_t = \hat{y}\frac{n-\beta}{n\beta - 1}E_t$$

$$c\overline{D}_r = \hat{x}c\epsilon_o E_r$$

$$c\overline{D}_t = \hat{x}\frac{1}{c\mu_o}\left[p - l\frac{ck_t}{\omega_t}\right]E_t$$

$$= \hat{x}\frac{1}{c\mu_o}\frac{n(n-\beta)}{1-n\beta}E_t$$

$$\overline{H}_r = -\hat{y}\frac{1}{c\mu_o}E_r$$

$$\overline{H}_t = \hat{y}\frac{1}{c\mu_o}\left[l + q\frac{ck_t}{\omega_t}\right]E_t = -\hat{y}\frac{1}{c\mu_o}nE_t$$

where $n = c\sqrt{\mu_o\epsilon_t}$ and $\beta = \nu/c$.

(b) To calculate amplitudes for the reflected and transmitted waves, apply the moving boundary conditions (1.5.19–1.5.20), which require that $E_x - \beta cB_y$ and $H_y - \beta cD_x$ be continuous across the boundary. Show that the reflection and transmission coefficients are

$$R = -\frac{1+\beta}{1-\beta}\frac{n-1}{n+1}$$

$$T = \frac{1-n\beta}{1-\beta}\frac{2}{n+1}$$

The reflectivity and transmissivity are

$$r = \frac{\hat{z}\cdot(\overline{E}_r \times \overline{H}_r)}{-\hat{z}\cdot(\overline{E}_i \times \overline{H}_i)} = |R|^2$$

$$t = \frac{-\hat{z}\cdot(\overline{E}_t \times \overline{H}_t)}{-\hat{z}\cdot(\overline{E}_i \times \overline{H}_i)} = n|T|^2$$

Show that $r + t \neq 1$ when $\beta \neq 0$. Is power conservation being violated?

(c) To answer this question [Daly and Gruenberg, 1967], conceive a cylinder of unit cross-section erected across the boundary with its axis parallel to the z axis and containing a portion of the interface. Show that the sum of the incident, reflected, and transmitted waves gives the total time-average electromagnetic power flow into the cylinder:

$$< P_{elec} >= \frac{1}{2c\mu_o}(\overline{E}^2 - E_r^2 - nE_t^2) = 4cU_0\beta\frac{(n-1)(1-n\beta)}{(1-\beta)^2(1+n)}$$

where

$$U_0 = \frac{1}{2c\mu_o}|E_0|^2$$

Inside the cylinder, there is an increase in the time-average electromagnetic energy, $Re(\overline{E}\cdot\overline{D}^{*}+\overline{H}\cdot\overline{B}^{*})/4$ as the moving dielectric occupies more free space. The rate of this increase in the stored energy is given by the velocity times the difference between the electromagnetic energy in the dielectric and that in the vacuum. The rate of increase of stored energy as the dielectric occupies more free space is given by the velocity times the difference in the volume densities of stored energy in the vacuum and in the dielectric. Show that

$$< P_{stored} > = \frac{\beta}{2c\mu_o}\left\{n\left[\frac{n-\beta}{1-n\beta}\right]E_t^2 - E_r^2 - E^2\right\}$$
$$= 2cU_0\frac{(n-1)(1-2n\beta+\beta^2)}{(1-\beta^2)(1+n)}$$

When the medium is stationary, $< P_{elec} >=< P_{stored} >= 0$. When the medium is in motion, this equality does not hold because mechanical power is required to keep the dielectric moving at constant velocity. The rate at which mechanical work has to be supplied to the system is given by the difference between $< P_{stored} >$ and $< P_{elec} >$. Show that

$$< P_{mech} >=< P_{stored} > - < P_{elec} >= -\frac{2cU_0\beta(n-1)(1+\beta)}{(1-\beta)(1+n)}$$

The negative sign indicates that mechanical work has been done to the system. The force per unit area acting on the dielectric medium is obtained from $\overline{F}\cdot\overline{v} =< P_{mech} >$. Thus

$$\overline{F}_{mech} = -\frac{\hat{z}2U_0(n-1)(1+\beta)}{(1-\beta)(1+n)}$$

This mechanical force is needed to maintain the medium at constant velocity. We note that the force is in the negative $\hat{z}$ direction; this means mechanical force must be applied to stop the medium from accelerating toward the wave. The electromagnetic force $\overline{F}_{elec}$ exerted on the medium by the wave is equal to the negative of $\overline{F}_{mech}$, as required by the basic laws of mechanics.

(d) Double-check this assertion by using the conservation theorem to calculate the electromagnetic force $\overline{F}_{elec}$. When the force density

is integrated over the volume of the cylinder, Show that the force per unit area acting on the surface is

$$\overline{F}_{elec} = \hat{z}\left(< T_{zz} >_i + < T_{zz} >_r - < T_{zz} >_t\right) + v\left(< \overline{G} >_i + < \overline{G} >_r - < \overline{G} >_t\right)$$
$$= \frac{\hat{z}2U_0(n-1)(1+\beta)}{(1-\beta)(1+n)}$$

Clearly, $\overline{F}_{elec}$ and $\overline{F}_{mech}$ are indeed in opposite directions. Thus the radiation pressure exerted on a dielectric half-space by a plane wave at normal incidence results in a force attracting the medium toward the wave. This force is there whether the dielectric is stationary or in motion. A mechanical force counterbalancing $\overline{F}_{elec}$ is needed either to keep the dielectric medium stationary or to maintain its constant velocity when it is in motion.

(e) Show that for a perfect conductor,

$$< P_{mech} > = \frac{1}{2c\mu_o}\left[-(E^2 + E_r^2) - (E^2 - E_r^2)\right]$$
$$= \frac{2cU_0\beta(1+\beta)}{(1-\beta)}$$

It follows that $\overline{F}_{mech} = \hat{z}2U_0(1+\beta)/(1-\beta)$. Note in particular the sign of $\overline{F}_{mech}$ which is now in the positive $\hat{z}$ direction, demonstrating that the wave is exerting a force to push the conductor away. Thus the electromagnetic force is attractive when the medium is a dielectric, and repulsive when the medium is a perfect conductor.

Problem P7.15

Consider reflection and transmission by a moving uniaxial media with optic axis along the $\hat{z}$ direction. Assume that the plane of incidence is parallel to the direction of motion.

(a) Show that the dispersion relation is

$$k_x^2 + \frac{\mu_z}{\mu}(k_z - \omega\xi)^2 = \omega^2\mu_z\epsilon$$

for TE waves, and

$$k_x^2 + \frac{\epsilon_z}{\epsilon}(k_z - \omega\xi)^2 = \omega^2\mu\epsilon_z$$

for TM waves. Note that the TE and TM waves correspond to the Type I and Type II waves discussed in Chapter II.

(b) Assume that the incident region 0 is also a moving medium. Total reflection occurs when the transmitted wave is evanescent. Find critical angles by requiring that $k_{tx}^2 \leq 0$. The condition for total reflection gives

$$(\omega - \chi_t k_z)^2 - \kappa_t \nu_t k_z^2 \leq 0$$

Consider moving isotropic media and approximate $\kappa\nu \approx c^2/n^2$, $\kappa_t\nu_t \approx c^2/n_t^2$, $\chi \approx c\beta(1-1/n^2)$, and $\chi_t \approx c\beta_t(1-1/n_t^2)$ keeping only terms of the first order in β, . For non-relativistic velocities, $\omega = \chi k_z + ck/n$. Show that

$$\sin\theta \geq \frac{n_t/n}{1 + n_t\beta_t(1-1/n_t^2) - n_t\beta(1-1/n^2)} \approx n_t\beta(1-1/n^2)$$

The last two terms are due to the motion of the media in regions 0 and t. The faster the medium in region t moves, the smaller the critical angle. Motion of the medium in region 0 induces opposite effects.

(c) Show that the reflection coefficients due to the stratified moving media take the same forms as those due to the stationary ones. For two isotropic media moving with non-relativistic velocities, $\omega \approx \chi k_z + ck/n$. Calculate Brewster's angle by setting the Fresnel reflection coefficients equal to zero.

Problem P7.16

The space–time transformation from an integral frame to the same frame rotating with angular velocity Ω around the z axis is given non-relativistically by

$$\begin{aligned} x &= x'\cos\Omega t' + y'\sin\Omega t' \\ y &= -x'\sin\Omega t' + y'\cos\Omega t' \\ z &= z' \\ t &= t' \end{aligned}$$

Show that, in order to preserve the form invariance of Maxwell's equations, the current charge densities and the field vectors must transform

as follows:

$$\begin{aligned}
\overline{J} &= \overline{J}' - \rho'\overline{\Omega} \times \overline{r}' \\
\rho &= \rho' \\
\overline{E} &= \overline{E}' + (\overline{\Omega} \times \overline{r}') \times \overline{B}' \\
\overline{H} &= \overline{H}' - (\overline{\Omega} \times \overline{r}') \times \overline{D}' \\
D &= D' \\
B &= B'
\end{aligned}$$

where $\overline{\Omega} = \hat{z}\Omega$, and $\overline{r}$ is the position vector. Find the constitutive relations for an isotropic medium as viewed from the rotating frame, under the assumption that the velocity $\overline{\Omega} \times \overline{r} = \overline{\Omega} \times \overline{r}'$ is small.

Problem P7.17

Show that the homogeneous Lorentz transformations (LT's) form a group by demonstrating that (a) two successive LT's are equivalent to another LT (closure), (b) the associative law applies to successive LT's, (c) the identity of the group is no transformation at all, and (d) for each LT there exists an inverse LT.

Problem P7.18

Show that pure Lorentz transformations violate the closure postulate of a group.

REFERENCES

Abramowitz, M. and I. A. Stegun, *Handbook of Mathematical Functions*, Dover Publications, New York, 1965.

Adams, A. T., *Electromagnetics for Engineers*, Renold Press, New York, 1971.

Agarwal, G. S., "Interaction of electromagnetic waves at rough dielectric surfaces," *Phys. Rev. B*, **15**, 2371–2383, 1977

Ali, S. M., W. C. Chew, and J. A. Kong, "Vector Hankel transform analysis of annular-ring microstrip antenna," *IEEE Trans. Antennas Propagat.*, **AP-30**, 637–644, 1982.

Anderson, J. L., *Principles of Relativity Physics*, Academic Press, New York, 1967.

Anderson, J. L. and J. W. Ryon, "Electromagnetic radiation in accelerated system," *Phys. Rev.*, **181**, 1765–1775, 1969.

Appel-Hansen, J., "The loop antenna with director arrays of loops and rods," *IEEE Trans. Antennas Propagat.*, **AP-20**, 516–517, 1972.

Arnaud, J. A. and A. A. M. Saleh, "Guidance of surface waves by multilayer coatings," *Appl. Optics*, **13**, 2343–2345, 1974.

Astrov, D. N., "The magnetoelectric effect in antiferromagnetics," *Zh. Eksp. Teor. Fiz.*, **38**, 984–985, 1960.

Bahar, E., "Propagation of radio waves over a nonuniform layered medium," *Radio Sci.*, **5**, 1069–1076, 1970.

Balanis, C. A., *Antenna Theory: Analysis and Design*, Harper & Row, New York, 1982.

Bannister, P. R., "The image theory for electromagnetic fields of a horizontal electric dipole in the presence of a conducting half space," *Radio Sci.*, **17**, 1095–1102, 1982.

Baños, A., Jr., *Dipole Radiation in the Presence of a Conducting Half-Space*, Pergamon Press, New York, 1966.

Barabanenkov, Y. N., Y. A. Kravtsov, S. M. Rytov, and V. I. Tatarskii, "Status of the theory of propagation of wave in a randomly inhomoge-

neous medium," *Soviet Phys. Usp.*, **13**, 551–580, 1971.

Barrick, D. E. and W. H. Peake, "A review of scattering from surfaces with different roughness scales," *Radio Sci.*, **3**, 865–868, 1968.

Barut, A. O., *Electrodynamics and Classical Theory of Fields and Particles*, MacMillan, New York, 1964.

Beckmann, P. and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces*, Pergamon Press, New York, 1963.

Bennett, C. L. and G. F. Ross, "Time-domain electromagnetics and its applications," *Proc. IEEE*, **66**, 299–318, 1978.

Bergmann, P. G., *Introduction to the Theory of Relativity*, Prentice-Hall, Englewood Cliffs, NJ, 1942.

Besieris, I. M. and W. E. Kohler, "Two-frequency radiative transfer equation for statistically inhomogeneous and anisotropic absorptive medium," in *Multiple Scattering and Waves in Random Media*, edited by P. L. Chow, W. E. Kohler, and G. C. Papanicolaou, North-Holland Publishing Company, New York, 1981.

Birss, R. R., "Macroscopic symmetry in space–time," *Rep. Prog. Phys.*, **26**, 307–310, 1963.

Birss, R. R. and R. G. Shrubsall, "The propagation of EM Waves in magnetoelectric crystals," *Phil. Mag.*, **15**, 687–700, 1967.

Bjorken, J. D. and S. D. Drell, *Relativistic Quantum Fields*, McGraw-Hill, New York, 1965.

Blanchard, A. J. and J. W. Rouse, Jr., "Depolarization of electromagnetic waves scattered from an inhomogeneous half space bounded by a rough surface", *Radio Sci.*, **15**, 773–780, 1980.

Boerner, W. M., M. B. El-Arini, C. Y. Chan, and P. M. Mastoris, "Polarization dependence in electromagnetic inverse problems," *IEEE Trans. Antennas Propagat.*, **AP-29**, 262–271, 1981.

Bolotovskii, B. M. and A. N. Lebedev, "On threshold phenomena in classical electrodynamics," *Soviet Phys. JETP*, **26**, 784, 1968.

Booker, H. G., *Energy in Electromagnetism*, Peter Peregrinus, New York, 1982.

Born, M. and E. Wolf, *Principles of Optics*, Pergamon Press, New York, 1970.

Botros, A. Z. and S. F. Mahmoud, "The transient fields of simple radiators from the point of view of remote sensing of the ground subsurface," *Radio Sci.*, **13**, 379–389, 1978.

Bouwkamp, C. J. and H. B. G. Casimir, "On multipole expansions in the theory of electromagnetic radiation," *Physica*, **20**, 539–554, 1954.

Bowman, J. J., T. B. A. Senior, and P. L. E. Uslenghi, *Electromagnetic and Acoustic Scattering by Simple Shapes*, North Holland, Amsterdam, 1969.

Boyd, G. D. and J. P. Gordon, "Confocal multimode resonator for millimeter through optical wavelength masers," *Bell System Tech. J.*, **40**, 489–508, 1961.

Brekhovskikh, L. M., *Waves in Layered Media*, Academic Press, New York, 1960.

Brevik, I. and B. Lautrap, "Quantum electrodynamics in material media," *K. Dan Vidensk, Selsk, Mat. Fys. Medd.*, **38**, 3–36, 1970.

Brown, G. S.,"The validity of shadowing corrections in rough surface scattering," *Radio Sci.*, **19**, 1461–1468, 1984.

Burke, H.-H. K. and T. J. Schmugge, "Effects of varying soil moisture contents and vegetation canopies on microwave emissions," *IEEE Trans. Geosci. Remote Sensing*, **GE-20**, 268–274, 1982.

Burke, W. J., T. Schmugge, and J. F. Paris, "Comparison of 2.8 and 1 cm microwave radiometer observations over soils with emission model calculations," *J. Geophys. Res.*, **84**, 287–294, 1979.

Carniglia, C. D. and L. Mandel, "Quantization of evanescent electromagnetic waves," *Phys. Rev. D*, **3**, 280–296, 1971.

Čerenkov, P. A., "Visible radiation produced by electrons moving in a medium with velocities exceeding that of light," *Phys. Rev.*, **52**, 378–379, 1937.

Chang, D. C., E. F. Kuester, and A. R. Mahmad, "Geometrical theory of a one-dimensional microstrip resonator: the effect of topside charges and currents," *Radio Sci.*, **20**, 819–826, 1985.

Chang, S. K. and K. K. Mei, "Application of the unimoment method to electromagnetic scattering of dielectric cylinders," *IEEE Trans. Antennas Propagat.*, **AP-24**, 35–42, 1976.

Chang, T. C., P. Gloersen, T. Schmugge, T. T. Wilheit, and H. J.

Zwally, "Microwave emission from snow and glacier ice," *J. Glaciology*, **16**, 23–29, 1976.

Chari, M. V. K. and P. P. Silvester, *Finite Elements in Electric and Magnetic Field Problems*, John Wiley & Sons, New York, 1980.

Chawla, B. R. and H. Unz, *Electromagnetic Waves in Moving Magneto-Plasmas*, University Press, Lawrence, Kansas, 1971.

Chen, H. C. and D. K. Cheng, "Dyadic Green's function for anisotropic and compressible media," *Proc. IEEE*, **53**, 1268–1269, 1965.

Cheng, D. K., *Field and Wave Electromagnetics*, Addison-Wesley, MA, 1983.

Cheng, D. K. and B. J. Strait, "An unusually simple method for side-lobe reduction," *IEEE Trans. Antennas Propagat.*, **AP-11**, 375–376, 1963.

Chew, W. C. and J. A. Kong, "Asymptotic formula for the capacitance of two oppositely charged discs," *Math. Proc. Camb. Phil. Soc.*, **89**, 373–384, 1981.

Chu, L. J. and J. A. Stratton, "Forced oscillations of a prolate spheroid," *J. Appl. Phys.*, **12**, 241–248, 1941.

Chu, R. S., J. A. Kong, and T. Tamir, "Diffraction of Gaussian beams by a periodically-modulated layer," *J. Opt. Soc. Am.*, **67**, 1555–1561, 1977.

Chu, R. S. and J. A. Kong, "Diffraction of optical beams with arbitrary profiles by a periodically-modulated layer," *J. Opt. Soc. Am.*, **70**, 1–6, 1980.

Chu, T. S. and D. C. Hogg, "Effects of precipitation on propagation at 0.63, 3.5 and 10.6 microns," *Bell System Tech. J.*, **47**, 723–759, 1968.

Chuang, S. L. and J. A. Kong, "Wave scattering and guidance by dielectric waveguides with periodic surface," *J. Opt. Soc. Am.*, **73**, 669–679, 1983.

Chuang, S. L., J. A. Kong, and L. Tsang, "Radiative transfer theory for passive microwave remote sensing of a two-layer random medium with cylindrical structures," *J. Appl. Phys.*, **51**, 5588–5593, 1980.

Chuang, S. L., L. Tsang, J. A. Kong, and W. C. Chew, "The equivalence of the electric and magnetic surface current approaches in microstrip antenna studies," *IEEE Trans. Antennas Propagat.*, **AP-28**,

569–571, 1980.

Clemmow, P. C., *The Plane Wave Spectrum Representation of Electromagnetic Fields*, Pergamon Press, New York, 1966.

Colbeck, S., "The geometry and permittivity of snow at high frequencies," *J. Appl. Phys.*, **53**, 4495–4500, 1982.

Cole, J. D., *Perturbation Methods in Applied Mathematics*, Blaisdell, Waltham, Mass., 1968.

Collier, J. R. and C. T. Tai, "Guided waves in moving media," *IEEE Trans. Microwave Theory Tech.*, **MTT-13**, 441–445, 1965.

Collin, R. E., *Field Theory of Guided Waves*, McGraw-Hill, New York, 1960.

Collin, R. E. and F. J. Zucker, *Antenna Theory*, McGraw-Hill, New York, 1969.

Compton, R. T., Jr., "The time-dependent Green's function for electromagnetic waves in moving simple media," *J. Math. Phys.*, **7**, 2145–2152, 1966.

Corson, D. and P. Lorrain, *Electromagnetic Waves and Fields*, Freeman, San Francisco, 1962.

Costen, R. C. and D. Adamson, "Three-dimensional derivation of the electrodynamic jump conditions and momentum-energy laws at a moving boundary," *Proc. IEEE*, **53**, 1181–1196, 1965.

Cox, D. C., "Depolarization of radio waves by atmospheric hydrometeors in earth-space paths: A review," *Radio Sci.*, **16**, 781–812, 1981.

Crane, R. K., "Propagation phenomena affecting satellite communication systems operating in the centimeter and millimeter wavelength bands," *Proc. IEEE*, **59**, 173–188, 1971.

Daly, P. and H. Gruenberg, "Energy relations for plane waves reflected from moving media," *J. Appl. Phys.*, **38**, 4486–4489, 1967.

Dashen, R., "Path integrals for waves in random media," *J. Math. Phys.*, **20**, 894–920, 1979.

Davenport, W. B., Jr. and W. L. Root, *An Introduction to the Theory of Random Signals and Noise*, McGraw-Hill, New York, 1958.

Deringin, L. N., "The reflection of a longitudinally polarized plane wave from a surface of rectangular corrugations," *Radio Tekhnol.*, **15**,

9–16, 1960.

DeGroot, S. R. and L. G. Suttorp, *Foundations of Electrodynamics*, North-Holland, Amsterdam, 1972.

DeSanto, J. A., "Green's function for electromagnetic scattering from a random rough surface," *J. Math. Phys.*, **15**, 283–288, 1974.

DeVries, H., "Rotatory power and other properties of certain liquid crystals," *Acta Crystallogs.*, **4**, 219–226, 1951.

Deirmendjian, D., *Electromagnetic Scattering on Spherical Polydispersions*, Elsevier, New York, 1969.

Dirac, P. A. M., *Principles of Quantum Mechanics*, 4th ed., Oxford University Press, London, 1958.

Dirac, P. A. M., *Lectures on Quantum Field Theory*, Yeshiva University, New York, 1966.

Djermakoye B. and J. A. Kong, "Radiative transfer theory for the remote sensing of layered media," *J. Appl. Phys.*, **50**, 6600–6604, 1979.

Du, L. J. and R. T. Compton, Jr., "Cutoff phenomena for guided waves in moving media," *IEEE Trans. Microwave Theory Tech.*, **MTT-14**, 358–363, 1966.

Dzyaloshinskii, I. E., "On the magnetoelectrical effect in antiferromagnets," *Soviet Phys. JETP*, **10**, 628–269, 1960.

Elachi, C., "Dipole antenna in space–time periodic media," *IEEE Trans. Antennas Propagat.*, **AP-20**, 280–287, 1972.

Elliott, R. S., *Antenna Theory and Design*, Prentice-Hall, New York, 1981.

England, A. W., "Thermal microwave emission from a scattering layer," *J. Geophys. Res.*, **80**, 4484–4496, 1975.

Evans, S., "Dielectric properties of ice and snow — A review," *J. Glaciology*, 773–792, 1965.

Fano, F. M., L. J. Chu, and R. B. Adler, *Electromagnetic Fields, Energy, and Forces*, Wiley, New York and M.I.T. Press, Cambridge, Mass., 1960.

Fante, R. L., "Relationship between radiative-transport theory and Maxwell's equations in dielectric media," *J. Opt. Soc. Am.*, **71**, 460–468, 1981.

Felsen, L. B. and N. Marcuwitz, *Radiation and Scattering of Electromagnetic Waves*, Prentice-Hall, Englewood Cliffs, New Jersey, 1973.

Feynman, R. P. and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965.

Fikioris, J. G. and P. C. Waterman, "Multiple scattering of waves: II. Hole corrections in the scalar case," *J. Math. Phys.*, **5**, 1413–1420, 1964

Foldy, L. L., "The multiple scattering of waves," *Phys. Rev.*, **67**, 107–119, 1945.

Frank, I. and Ig. Tamm, "Coherent visible radiation of fast electrons passing through matter," *Compt. Rend. (Dokl.)*, **14**, 109–114, 1937.

Frankl, D. R., *Electromagnetic Theory*, Prentice-Hall, New Jersey, 1986.

Frisch, V., "Wave propagation in random medium," in *Probabilistic Methods in Applied Mathematics*, **1**, edited by Bharuch-Reid, Academic Press, 1968.

Fuchs, R., "Wave propagation in a magnetoelectric medium," *Phil. Mag.*, **11**, 647, 1965.

Fulton, R. L., "Macroscopic quantum electrodynamics," *Jour. Chem. Phys.*, **50**, 3355–3377, 1969.

Fung, A. K. and H. L. Chan, "Backscattering of waves by composite rough surfaces," *IEEE Trans. Antennas Propagat.*, **AP-17**, 590–597, 1969.

Fung, A. K. and R. K. Moore, "The correlation function in Kirchhoff's method of solution of scattering of waves from statistically rough surfaces," *J. Geophys. Res.*, **71**, 2939–2943, 1966.

Furutsu, K., "Multiple scattering of waves in a medium of randomly distributed particles and derivation of the transport equation," *Radio Sci.*, **10**, 29–44, 1975.

Furutsu, K., "Statistical theory of scattering and propagation over a random surface," *Proc. IEE*, **130**, 601–622, 1983.

Garcia, N., V. Celli, N. R. Hill, and N. Cabrera, "Ill-conditioned matrices in the scattering of waves from hard corrugated surfaces," *Phys. Rev. B*, **18**, 5184–5189, 1978.

Ginzburg, V. L., "Radiation of an electron moving in a crystal with a

constant velocity exceeding that of light," *J. Phys.*, **3**, 101–106, 1940.

Ginzburg, V. L., *Propagation of Electomagnetic Waves in Plasma*, Addison-Wesley, Reading, Mass., 1964.

Goldstein, H., *Classical Mechanics*, Addison-Wesley, Reading, Mass., 1950.

Goodman, J. W., *Introduction to Fourier Optics*, McGraw-Hill, New York, 1968.

Goos, Von F. and H. Hänchen, "Ein neuer und fundamentaler Versuch zur Totalreflexion," *Ann. der Phys.*, **6**, 333–346, 1947.

Gradshteyn, I. S. and I. M. Ryzhik, *Table of Integrals, Series and Products*, Academic Press, New York, 1965.

Grant, I. S. and W. R. Phillips, *Electromagnetism*, John Wiley & Sons, New York, 1975.

Gray, K.G. and S.A. Bowhill, "Transient response of stratified media: multiple scattering integral and differential equation for an impulsive incident plane wave," *Radio Sci.*, **9**, 57–62, 1974.

Gray, E. P., R. W. Hart, and R. A. Farrell, "An application of a variational principle for scattering by random rough surfaces," *Radio Sci.*, **13**, 333–343, 1978.

Gruodis, A.J. and C. S. Chang, "Coupled lossy transmission line characterization and simulation," *IBM J. Res. Develop.*, **25**, 1981.

Gurvich, A. S., V. L. Kalinin, and D. T. Matveyer, "Influence of the internal structure of glaciers on their thermal radio emission," *Atm. Oceanic Phys. USSR*, **9**, 713–717, 1973.

Gurvich, A. S. and V. I. Tatarskii, "Coherent and intensity fluctuations of light in the turbulent atmosphere," *Radio Sci.*, **10**, 3–14, 1975.

Habashy, T. M., J. A. Kong, and L. Tsang, "Quasi-static electromagnetic fields due to dipole antennas in bounded conducting media," *IEEE Trans. Geosci. Remote Sensing*, **GE-23**, 325–333, 1985.

Hansen, P. M. and C. T. Tai, "Radiation from sources in the presence of a flat Earth," *IEEE Trans. Antennas Propagat.*, **AP-18**, 423–424, 1970.

Hansen, W. W. and J. R. Woodyard, "A new principle in directional antenna design," *Proc. IRE*, **26**, 333–345, 1938.

Harrington, R. F., *Time-Harmonic Electromagnetic Fields*, McGraw-Hill, New York, 1961.

Harrington, R. F., *Field Computation by Moment Methods*, MacMillan, New York, 1968.

Harrington, R. F. and A. T. Villeneuve, "Reciprocity relationships for gyrotropic media," *IRE Trans. Microwave Theory Tech.*, **MTT-6**, 308–310, 1958.

Heitler, W., *The Quantum Theory of Radiation*, 3rd ed., Oxford University Press, London, 1966.

Hertz, H., *Electric Waves*, MacMillan and Co., New York, 1893.

Hill, D. A. and J. R. Wait "Excitation of the Zenneck surface wave by a vertical aperture," *Radio Sci.*, **13**, 969–977, 1978.

Hill, E. L., "Hamilton's principle and the conservation theorems of mathematical physics," *Rev. Mod. Phys.*, **23**, 253–260, 1951.

Holt, A. R., N. K. Uzunoglu, and B. G. Evans, "An integral equation solution to scattering of electromagnetic radiation by dielectric spheroids and ellipsoids," *IEEE Trans. Antennas Propagat.*, **AP-26**, 706–712, 1978.

de Hoop, A. T. and H. J. Frankena, "Radiation of pulses generated by a vertical electric dipole above a plane non-conducting earth," *Appl. Sci. Res.*, **B8**, 369–377, 1960.

Hu, C. and J. R. Whinnery, "Field-realigned nematic-liquid-crystal optical waveguides," *IEEE J. Quantum Electron.*, **QE-10**, 556–562, 1974.

Huang, H. C., *Coupled Mode Theory*, VNU Science Press, The Netherlands, 1984.

Hutley, M. C. and V. M. Bird, "A detailed experimental study of the anomalies of a sinusoidal diffraction grating," *Opt. Acta.*, **20**, 771–782, 1973.

Indenbom, V. L., "Irreducible representations of the magnetic groups and allowance for magnetic symmetry," *Soviet Phys. Crystallogr.*, **5**, 493, 1960.

Ishimaru, A., *Wave Propagation and Scattering in Random Media*, Academic Press, New York, 1978.

Ishimaru, A., D. Lesselier, and C. Yeh, "Multiple scattering calculations for nonspherical particles based on the vector radiative transfer theory," *Radio Sci.*, **19**, 1356–1366, 1984.

Ishimaru, A., R. Woo, J. W. Armstrong, and D. C. Backman, "Multiple scattering calculation of rain effects," *Radio Sci.*, **17**, 1425–1433, 1982.

Ito, S. and S. Adachi, "Multiple scattering effect on backscattering from a random medium", *IEEE Trans. Antennas Propagat.*, **AP-25**, 205–208, 1977.

Itoh, T., "Generalized spectral domain method for multiconductor printed lines and its application to tunable suspended microstrips," *IEEE Trans. Microwave Theory Tech.*, **MTT-26**, 983–987, 1978.

Jackson, J. D., *Classical Electrodynamics*, Wiley, New York, 1962.

James, G. L., *Geometrical Theory of Diffraction for Electromagnetic Waves* Peter Peregrinus, England, 1976.

James, J. R., P. S. Hall, and C. Wood, *Microstrip Antenna*, Peter Peregrinus, New York, 1981.

Jao, J. K., "Amplitude distribution of composite terrain radar clutter and the K-distribution," *IEEE Trans. Antennas Propagat.*, **AP-32**, 1049–1062, 1984.

Jauch, J. M. and K. M. Watson, "Phenomenological quantum electrodynamics," *Phys. Rev.*, **74**, 950–957, 1948.

Jauch, J. M. and K. M. Watson, "Phenomenological quantum electrodynamics. Part II: Interaction of the field with charges," *Phys. Rev.*, **75**, 1485–1493, 1948.

Jauch, J. M. and K. M. Watson, "Phenomenological quantum electrodynamics. Part III: Dispersion," *Phys. Rev.*, **75**, 1249–1261, 1949.

Jordan, A. K. and R. H. Lang, "Electromagnetic scattering patterns from sinusoidal surface," *Radio Sci.*, **14**, 1077–1088, 1979.

Jordan, E. C. and K. G. Balmain, *Electromagnetic Waves and Radiating Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1968.

Kapany, N. S. and J. J. Burke, *Optical Waveguides*, Academic Press, New York, 1972.

Katehi, P. B. and N. G. Alexopoulos, "On the modeling of electromagnetically coupled microstrip antennas — the printed strip dipole,"

IEEE Trans. Antennas Propagat., **AP-32**, 1179–1186, 1984.

Kim, H. T., N. Wang, and D. L. Moffatt, "K-pulse for a thin circular loop," *IEEE Trans. Antennas Propagat.*, **AP-33**, 1403–1407, 1985.

King, R. W. P. and C. W. H. Harrison, Jr., *Antennas and Waves*, M.I.T. Press, Cambridge, Mass., 1969.

King, R. W. P., R. B. Mack, and S. S. Sandler, *Arrays of Cylindrical Dipoles*, Cambridge Univ. Press, New York, 1968.

Klauder, J. R., and E. C. G. Sudarshan, *Fundamentals of Quantum Optics*, Benjamin, New York, 1968.

Klein, W. R. and B. D. Cook, "Unified approach to ultrasonic light diffraction," *IEEE Trans. Sonics Ultrason.*, **SU-14**, 123–134, 1967.

Kogelnik, H., "Coupled wave theory for thick hologram gratings," *Bell System Tech. J.*, **48**, 2909–2947, 1969.

Kogelnik, H. and C. V. Shank, "Simulated emission in a periodic structure," *Appl. Phys. Lett.*, **18**, 152–154, 1971.

Kogelnik, H. and C. V. Shank, "Coupled-wave theory of distributed feedback lasers," *J. Appl. Phys.*, **43**, 2327–2335, 1972.

Kohler, W. E. and G. C. Papanicolaou, "Some applications of the coherent potential approximation," in *Multiple Scattering and Waves in Random Media*, edited by P. L. Chow, W. E. Kohler, and G. C. Papanicolaou, 199–223, North-Holland Publishing Company, New York, 1981.

Kong, J. A., "Theorems of bianisotropic media," *Proc. IEEE*, **60**, 1036–1046, 1972.

Kong, J. A., "Dispersion analysis of reflection and transmission by a plane boundary — a graphical approach," *Am. J. Phys.*, **43**, 73–76, 1975.

Kong, J. A., *Theory of Electromagnetic Waves*, Wiley-Interscience, New York, 1975.

Kong, J. A., "Second-order coupled mode equations for spatially periodic media," *J. Opt. Soc. Am.*, **67**, 825–829, 1977.

Kong, J. A., *Research Topics in Electromagnetic Wave Theory*, Wiley-Interscience, New York 1981.

Kraus, J. D., *Electromagnetics*, McGraw-Hill, New York, 1984.

Kravtsov, Y. A. and A. I. Saichev, "Effects of double passage of waves in randomly inhomogeneous media," *Soviet Phys. Usp.*, **25**, 494–508, 1982.

Kritikos, H. N., D. L. Jaggard, and D. B. Ge, "Numeric reconstruction of smooth dielectric profiles," *Proc. IEEE*, **70**, 295–297, 1982.

Kritikos, H. N. and J. Shiue, "Microwave remote sensing from orbit," *IEEE Spectrum*, 34–41, 1979.

Lanczos, C., *Applied Analysis*, Prentice-Hall, NJ, 1956.

Landau, L. D. and E. M. Lifshitz, *Electrodynamics of Continuous Media*, Addison-Wesley, Reading, Mass., 1960.

Landau, L. D. and E. M. Lifshitz, *Mechanics*, 2nd ed., Addison-Wesley, Reading, Mass., 1960.

Landau, L. D. and E. M. Lifshitz, *Classical Theory of Fields*, 3rd ed., Addison-Wesley, Reading, Mass., 1971.

Landau, L. D. and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory*, Addison-Wesley, Reading, Mass., 1958.

Lax, M., "Wave propagation and conductivity in random media," *Proc. SIAM-AMS*, **VI**, 35–95, 1973.

Leader, J. C., "Bidirectional scattering of electromagnetic waves from rough surfaces," *J. Appl. Phys.*, **42**, 4808–4816, 1971.

Lee, J. K. and J. A. Kong, "Dyadic Green's functions for layered anisotropic medium," *Electromagnetics*, **3**, 111–130, 1983.

Lee, K. F., *Principles of Antenna Theory*, John Wiley & Sons, New York, 1984.

Lee, K. S. H. and C. H. Papas, "Electromagnetic radiation in the presence of moving simple media," *J. Math. Phys.*, **5**, 1668–1672, 1964.

Lee, M. C., A. Das Gupta, J. A. Klobuchar, S. Basu, and S. Basu, "Depolarization of VHF geostationary satellite signals near the equatorial anomaly crests," *Radio Sci.*, **17**, 399–409, 1982.

Lee, S. W. and Y. T. Lo, "Reflection and transmission of electromagnetic waves by a moving uniaxially anisotropic medium," *J. Appl. Phys.*, **38**, 870–875, 1967.

Lindell, I. V. and A. H. Sihvola, "Dielectrically loaded corrugated waveguide: variational analysis of a nonstandard eigenproblem," *IEEE*

Trans. Microwave Theory Tech., **MTT-31**, 552–526, 1983.

Lindell, I. V., E. Alanen, and K. Mannersalo, "Exact image method for impedance computation of antennas above the ground," *IEEE Trans. Antennas Propagat.*, **AP-33**, 937–945, 1985.

Lo, Y. T., D. Solomon, and W. F. Richards, "Theory and experiment on microstrip antennas," *IEEE Trans. Antennas Propagat.*, **AP-27**, 137–145, 1979.

Lorentz, G., *Bernstein Polynominals*, Toronto University Press, Can., 1953.

Louisell, W., *Radiation and Noise in Quantum Electronics*, McGraw-Hill, New York, 1964.

Lynch, P. J. and R. J. Wagner, "Rough-surface scattering: shadowing, multiple scatter, and energy conservation," *J. Math. Phys.*, **11**, 3032–3042, 1970.

Ma, M. T., *Theory and Application of Antenna Arrays*, Wiley-Interscience, New York, 1974.

Marcuse, D., *Engineering Quantum Electrodynamics*, Van Nostrand-Reinhold, New York, 1970.

Marcuse, D., *Light Tranmission Optics*, Van Nostrand-Reinhold, New York, 1972.

Marion, J. B. and M. A. Heald, *Classical Electromagnetic Radiation*, 2nd ed., Academic Press, New York, 1980.

Mason, W. P., *Crystal Physics and Interaction Processes*, Academic Press, New York, 1966.

Matzler, C., E. Schanda, and W. Good, "Towards the definition of optimum sensor specifications for microwave remote sensing of snow," *IEEE Trans. Geosci. Remote Sensing*, **GE-20**, 57–66, 1982.

Maxwell, J. C., *A Treatise on Electricity and Magnetism*, Dover Publications, New York, 1954.

Maxwell-Garnett, J. C., *Trans. Roy. Soc. London*, **203**, 385, 1904.

McKenzie, J. F., "Electromagnetic waves in uniformly moving media," *Proc. Phys. Soc.*, **91**, 532–536, 1967.

Mei, K. K., "Unimoment method of solving antenna and scattering problems," *IEEE Trans. Antennas Propagat.*, **AP-22**, 760–766, 1974.

Melcher, J. R., *Continuum Mechanics*, MIT Press, Cambridge, Mass., 1981.

Mergelyan, O. S., "A point charge in a gyrotropic dielectric," *Soviet Phys. Tech. Phys.*, **12**, 594–597, 1967.

Mittra, R. and S. W. Lee, *Analytical Techniques in The Theory of Guided Waves*, Macmillan, New York, 1971.

Mo, T. C., "Theory of electrodynamics in media in noninertial frames and applications," *J. Math. Phys.*, **11**, 2589–2610, 1970.

Møller, C., *The Theory of Relativity*, Oxford University Press, London, 1966.

Morse, P. M. and H. Feshbach, *Methods of Theoretical Physics*, McGraw-Hill, New York, 1953.

O'Dell, T. H., "Magnetoelectrics — a new class of materials," *Electron. Power*, **11**, 266–267, 1965.

O'Dell, T. H., *The Electrodynamics of Magneto-Electric Media*, **11**: *Selected Topics in Solid State Physics*, E. P. Wohlforth (Ed.), North-Holland, Amsterdam, 1970.

Oguchi, T., "Electromagnetic wave propagation and scattering in rain and other hydrometeors," *Proc. IEEE*, **71**, 1029–1078, 1983.

Ogura, H., "Theory of waves in a homogeneous random medium," *Phys. Rev. A*, **11**, 942–956, 1975.

Oliner, A. A. and A. Hessel, "Guided waves on sinusoidally-modulated reactance surfaces," *IRE Trans. Antenna Propagat.*, S201–S208, 1959.

Onstott, R. G., R. K. Moore, and W. F. Weeks, "Surface-based scatterometer results of Arctic sea ice," *IEEE Trans. Geosci. Electron.*, **GE-17**, 78–85, 1979.

Panofsky, W. K. H. and M. Phillips, *Classical Electricity and Magnetism*, 2d ed., Addison-Wesley, Reading, Mass., 1962.

Papanicolaou, G. C. and J. B. Keller, "Stochastic differential equations with applications to random harmonic oscillators and wave propagation in random media," *SIAM J. Appl. Math.*, **21**, 287–305, 1971.

Papas, C. H., *Theory of Electromagnetic Wave Propagation*, McGraw-Hill, New York, 1965.

Parashar, S. K., R. M. Haralick, R. K. Moore, and A. W. Biggs, "Radar

scatterometer discrimination of sea-ice types," *IEEE Trans. Geosci. Electron.*, **GE-15**, 83–87, 1977.

Paris, D. T. and G. K. Hurd, *Basic Electromagnetic Theory*, McGraw-Hill, New York, 1969.

Peake, W. H., "Interaction of electromagnetic waves with some natural surfaces," *IEEE Trans. Antennas Propagat.*, **AP-7**, Special Supplement, S324–S329, 1959.

Peake, W. H., D. E. Barrick, A. K. Fung, and H. L. Chan, "Comments on 'Backscattering of waves by composite rough surfaces'," *IEEE Trans. Antennas Propagat.*, **AP-18**, 716–726, 1970.

Penfield, P., Jr. and H. A. Haus, *Electrodynamics of Moving Media*, M.I.T. Press, Cambridge, Mass., 1967.

Penfield, P. Jr., R. Spence, and S. Duinker, *Tellegen's Theorem and Electrical Networks*, M.I.T. Press, Cambridge, Mass, 1970.

Peng, S. T. and T. Tamir, "Directional blazing of waves guided by asymmetrical dielectric gratings," *Opt. Comm.*, **11**, 405-409, 1974.

Percus, J. K. and G. J. Yevick, "Analysis of classical statistical mechanics by means of collective coordinates," *Phys. Rev.*, **110**, 1–13, 1958.

Peterson, B. and S. Strom, "T matrix for electromagnetic scattering from an arbitrary number of scatterers and representation of E(3)," *Phy. Rev. D*, **8**, 3661–3678, 1973.

Plonus, M. A., *Applied Electromagnetics*, McGraw-Hill, New York, 1978.

Poh, S. Y., W. C. Chew, and J. A. Kong, "Approximate formulas for line capacitance and characteristic impedance of microstrip line," *IEEE Trans. Microwave Theory Tech.*, **MTT-29**, 135–142/1119, 1981.

Polder, D. and J. H. van Santen, "The effective permeability of mixtures of solids," *Physica*, **12**, 257–271, 1946.

Popovic, B. D., *Introductory Engineering Electromagnetic*, Addison-Wesley, Reading, Mass., 1971.

Post, E. J., "Electromagnetism and the principle of equivalence," *Ann. Phys.*, **70**, 507–515, 1972.

Pozar, D. M., D. H. Schaubert, and R. E. McIntosh, "The optimum

transient radiation from an arbitrary antenna," *IEEE Trans. Antennas Propagat.*, **AP-32**, 633–640, 1984.

Purcell, E. M., *Electricity and Magnetism*, McGraw-Hill, New York, 1963.

Rado, G. T., "Observation and possible mechanisms of magnetoelectric effects in a ferromagnet," *Phys. Rev. Lett.*, **13**, 355, 1964.

Raman, C. V. and N. S. N. Nath, "The diffraction of light by high frequency sound waves; I-V," *Proc. Ind. Acad. Sci.*, **2-3**, 1935–36.

Ramo, S., J. R. Whinnery, and T. Van Duzer, *Fields and Waves in Communication Electronics*, Pergamon Press, New York, 1970.

Rao, B. R. and T. T. Wu, "On the applicability of image theory in anisotropic media," *IEEE Trans. Antennas Propagat.*, **AP-13**, 814–815, 1965.

Rao, N. N., *Basic Electromagnetism with Applications*, Prentice-Hall, New Jersey, 1972.

Rao, S. M., D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas Propagat.*, **AP-30**, 409–418, 1982.

Read, F. H., *Electromagnetic Radiation*, John Wiley & Sons, New York, 1980.

Rice, S. O., "Reflection of EM waves by slightly rough surfaces," in *The Theory of Electromagnetic Waves*, edited by M. Kline, Interscience, New York, 1963.

Richards, W. F., S. E. Davidson, and S. A. Long, "Dual-band reactively loaded microstrip antenna," *IEEE Trans. Antennas Propagat.*, **AP-23**, 556–561, 1985.

Riley, J., William A. Davis, and L. M. Besieris, "The singularity expansion method and multiple scattering," *Radio Sci.*, **20**, 20–24, 1985.

Roetgen, W., C., "Ueber die durch Bewegung eines in homogenenen electrischen Felde befindlichen Dielectricums hervorgerufene electrodynamische Kraft," *Ann. Physik.*, **35**, 264–270, 1888.

Rohrlich, F., *Classical Charged Particles*, Addison-Wesley, Reading, Mass., 1965.

Rosenbaum, S., "The mean Green's function: a nonlinear approxima-

tion," *Radio Sci.*, **6**, 379–386, 1971.

Rubin, B. J. and H. L. Bertoni, "Scattering from a periodic array of conducting bars of finite surface resistance," *Radio Sci.*, **20**, 827–832, 1985.

Ruck, G. T., D. E. Barrick, W. D. Stuart, and C. K. Krichbaum, *Radar Cross-Section Handbook*, **1-2**, McGraw-Hill, New York, 1970.

Rumsey, V. H., "Reaction concept in electromagnetic theory," *Phys. Rev.*, **94**, 1483–1491; **95**, 1706, 1954.

Ryzhov, Y. A., V. V. Tamoikin, and V. I. Tatarskii, "Spatial dispersion of inhomogeneous media," *Soviet Phys. JETP*, **21**, 433–438, 1965.

Sarkar, T. K., M. F. Costa, C. H. I, and R. F. Harrington, "Electromagnetic transmission through mesh covered apertures and arrays of apertures in a conducting screen," *IEEE Trans. Antennas Propagat.*, **AP-32**, 908–913, 1984.

Saxton, J. A. and J. A. Lane, "Electrical properties of sea water," *Wireless Engineer*, 269–275, 1952.

Schelkunoff, S. A., *Electromagnetic Waves*, D. Van Nostrand, 1943.

Schelkunoff, S. A., *Advanced Antenna Theory*, Wiley, New York, 1952.

Schiff, L. I., *Quantum Mechanics*, McGraw-Hill, New York, 1968.

Schlomka, V. T., "Das Ohmsche Gesetz bei Bewegten Körpern," *Ann. Phys.*, **6**, 246–252, 1951.

Schmutzer, E. von, "Zur Relativistischer Elektrodynamik in Beliebigen Medien," *Ann. Phys. (Liepzig)*, **6**, 171–180, 1956.

Schwan, H. P., "Interaction of microwave and radio frequency radiation with biological systems," *IEEE Trans. Microwave Theory Tech.*, **MTT-19**, 146–152, 1971.

Scott, A. C. and F. Y. F. Chu, "Pulse saturation in a traveling wave parametric amplifier," *Proc. IEEE*, **62**, 1720–1721, 1974.

Sen, P. N., C. Scala, and M. H. Cohen, "A self-similar model for sedimentary rocks with applications to the dielectric constant of fused glass beads," *Geophysics*, **46**, 781–795, 1981.

Senior, T. B. A. and R. F. Goodrich, "Scattering by a sphere," *Proc. IEEE*, **111**, 907–916, 1964.

Seshadri, S. R., *Fundamentals of Transmission Lime and Electromagnetic Fields*, Addison-Wesley, Reading, Mass., 1971.

Sezginer, A. and J. A. Kong, "Transient response of line source excitation in cylindrical geometry," *Electromagnetics*, **4**, 35–54, 1984.

Shen, L. C. and J. A. Kong, *Electromagnetism*, Brooks/Cole, California, 1983.

Shin, R. T. and J. A. Kong, "Scattering of electromagnetic waves from a randomly perturbed quasiperiodic surface," *J. Appl. Phys.*, **56**, 10–21, 1984.

Shiozawa, T. and N. Kumagai, "Total reflection at the interface between relatively moving media," *Proc. IEEE*, **55**, 1243–1244, 1967.

Silver, S., *Microwave Antenna Theory and Design*, Dover Publications, New York, 1949.

Skolnik, M. I., *Introduction to Radar Systems*, McGraw-Hill, New York, 1980.

Smith, G. S., "Directive properties of antennas for transmission into a material half-space," *IEEE Trans. Antennas Propagat.*, **AP-32**, 232–246, 1984.

Snyder, A. W., "Understanding monomode optical fibers," *Proc. IEEE*, **69**, 6–13, 1981.

Sommerfeld, A., *Partial Differential Equations*, Academic Press, New York, 1962.

Sommerfeld, A., *Electrodynamics*, Academic Press, New York, 1949.

Sommerfeld, A., *Optics*, Academic Press, New York, 1949.

Staelin, D. H., "Passive remote sensing at microwave wavelengths," *Proc. IEEE*, **57**, 427–459, 1969.

Stogryn, A., "Electromagnetic scattering by random dielectric constant fluctuations in a bounded medium," *Radio Sci.*, **9**, 509–518, 1974.

Stratton, J. A., *Electromagnetic Theory*, McGraw-Hill, New York, 1941.

Stutzman, W. L. and G. A. Thiele, *Antenna Theory and Design*, John Wiley & Sons, New York, 1981.

Synge, J. L., *Relativity: The Special Theory*, North-Holland, Amster-

dam, 1965.

Tai, C. T., "A study of electrodynamics of moving media," *Proc. IEEE*, **52**, 685–689, 1964.

Tai, C. T., *Dyadic Green's Function in Electromagnetic Theory*, Intext Publishers, New York, 1971.

Tai, C. T., "On the eigenfunction expansion of dyadic Green's functions," *Proc. IEEE*, **61**, 480–481, 1973.

Tamoikin, V. V., "The average field in a medium having strong anisotropic inhomogeneities," *Radiophysics Quantum Electron.*, **14**, 228–233, 1971.

Tan, H. S., A. K. Fung, and H. Eom, "A second order renormalization theory for cross-polarized backscatter from a half space random medium," *Radio Sci.*, **15**, 1059–1065, 1980.

Tatarskii, V. I., "Propagation of electromagnetic waves in a medium with strong dielectric constant fluctuations," *Soviet Phys. JETP*, **19**, 946–953, 1964.

Tatarskii, V. I. and M. E. Gertsenshtein, "Propagation of waves in a medium with strong fluctuation of refractive index," *Soviet Phys. JETP*, **17**, 458–463, 1963.

Tatarskii, V. I., *Wave Propagation in a Turbulent Medium*, McGraw-Hill, New York, 1961.

Tatarskii, V. I., *The Effects of Turbulent Atmosphere on Wave Propagation*, National Tech. Information Service, 472, Springfield, VA, 1971.

Tellegen, B. D. H., "The Gyrator, a new electric network element," *Phillips Res. Rept.*, **3**, 81–101, 1948.

Tien, P. K. and R. Ulrich, "Theory of prism-film coupler and thin-film light guides," *J. Opt. Soc. Am.*, **60**, 1325–1337, 1970.

Tien, P. K., R. Ulrich, and R. J. Martin, "Optical second harmonic generation in form of coherent Čerenkov radiation from a thin-film waveguide," *Appl. Phys. Lett.*, **17**, 447–450, 1970.

von Tischer, M. and S. Hess, "Die Materialgleichungen in Beliebigen Medien," *Ann. Phys. (Leipzig)*, **3**, 113–121, 1969.

Tolman, R. C., *Relativity, Thermodynamics, and Cosmology*, Oxford University Press, London, 1966.

Towne, D. H., *Wave Phenomena*, Addison-Wesley, Reading, Mass., 1967.

Tsang, L., A. J. Blanchard, R. W. Newton, and J. A. Kong,"A simple relation between active and passive microwave remote sensing measurements of earth terrain," *IEEE Trans. Geosci. Remote Sensing*, **GE-20**, 482–485, 1982.

Tsang, L. and J. A. Kong, "Microwave remote sensing of a two-layer random medium," *IEEE Trans. Antennas Propagat.*, **AP-24**, 283–287, 1976.

Tsang, L., J. A. Kong, and R. T. Shin, *Theory of Microwave Remote Sensing*, Wiley-Interscience, New York, 1985.

Tseng, F. I. and D. K. Cheng, "A synthesis technique for linear arrays with wide-band elements," *Proc. IEEE*, **51**, 1679–1681, 1963.

Tuan, H. S. and C. H. Ou, "Scattering of a TM surface wave at a guide deformation," *J. Appl. Phys.* **44**, 5522–5525, 1973.

Twersky, V., "Coherent electromagnetic waves in pair-correlated random distributions of aligned scatterers," *J. Math. Phys.*, **19**, 215–230, 1978.

Tyras, G., *Radiation and Propagation of Electromagnetic Waves*, Academic Press, New York, 1969.

Van Bladel, J., *Electromagnetic Fields*, McGraw-Hill, New York, 1964.

Van de Hulst, H. C., *Light Scattering by Small Particles*, John Wiley & Sons, New York, 1957.

Van den Berg, P. M., "Diffraction theory of a reflection grating," *Appl. Sci. Res.*, **24**, 261–293, 1971.

Vant, M. R., R. O. Ramseier, and V. Makios, "The complex dielectric constant of sea ice at frequencies in the range 0.1–40 GHz," *J. Appl. Phys.*, 1264–1280, 1978.

Vezzetti, D. J. and J. B. Keller, "Refractive index, attenuation, dielectric constant and permeability of waves in a polarizable medium," *J. Math. Phys.*, **8**, 1861–1870, 1967.

Wait, J. R. (Ed.), *Electromagnetic Probing in Geophysics*, Golem Press, Boulder, CO, 1971.

Wait, J. R., *Electomagnetic Waves in Stratified Media*, 2nd ed., Perg-

amon Press, New York, 1970.

Wait, J. R., L. Thrane, and R. J. King, "The transient electric field response of an array of parallel wires on the Earth's surface," *IEEE Trans. Antennas Propagat.*, **AP-23**, 261, 1975.

Wang, S., M. L. Shah, and J. D. Crow, "Wave propagation in thin-film optical waveguides using gyrotropic and anisotropic materials as substrates," *IEEE J. Quantum Electron.*, **QE-8**, 212–216, 1972.

Wasylkiwskyj, W. and W. K. Kahn, "Element pattern bounds in uniform phased arrays," *IEEE Trans. Antennas Propagat.*, **AP-25**, 439–449, 1977.

Waterman, P. C. and R. Truell, "Multiple scattering of waves," *J. Math. Phys.*, **2**, 512–537, 1961.

Watson, G. N., *A Treatise on the Theory of Bessel Functions*, 2nd ed., Pergamon Press, New York, 1944.

Watson, J. G. and J. B. Keller, "Reflection, scattering, and absorption of acoustic waves by rough surfaces," *J. Acoust. Soc. Am.*, **74**, 1887–1894, 1983.

Weeks, W. T., L. L. Wu, M. F. McAllister, and A. Singh, "Resistive and inductive skin effect in rectangular conductors," *IBM J. Res. Develop.*, **23**, 652–660, 1979.

Wei, C., R. F. Harrington, J. R. Mautz, and T. K. Sarkar, "Multiconductor transmission lines in multilayered dielectric media," *IEEE Trans. Microwave Theory Tech.*, **32**, 439–449, 1984.

Weil, H. and C. M. Chu, "Scattering and absorption of electromagnetic radiation by thin dielectric discs," *Appl. Optics*, **15**, 1832–1836, 1976.

Wertheim, M. S., "Exact solution of the Percus-Yevick integral equation for hard spheres," *Phys. Rev. Lett.*, **20**, 321–323, 1963.

Whitman, G. M. and F. Schwering, "Scattering by periodic metal surfaces with sinusoidal height profiles — a theoretical approach," *IEEE Trans. Antennas Propagat.*, **AP-25**, 869–876, 1977.

Wilson, H. A., "On the electric effect of a rotating dielectric in a magnetic field," *Phil. Trans. Roy. Soc. London*, **204A**, 121–137, 1905.

Wilton, D. R, S. M. Rao, A. W. Glisson, D. H. Schaubert, O. M. Al-Bundak, and C. M. Butler, "Potential integrals for uniform and linear source distributions on polygonal and polyhedral domains," *IEEE*

Trans. Antennas Propagat., **AP-32**, 276–281, 1984.

de Wolf, D. A. "Propagation regimes for turbulent atmospheres," *Radio Sci.*, **10**, 53–57, 1975.

Wolff, E. A., *Antenna Analysis*, Wiley, New York, 1966.

Wu, T. T. and H. Lehmann, "Spreading of electromagnetic pulses," *J. Appl. Phys.*, **55**, 2064–2065, 1985.

Yaghjian, A. D., "Electric dyadic Green's functions in the source region," *Proc. IEEE*, **68**, 248–263, 1980.

Yamazawa, M., N. Inagaki, and T. Sekiguchi, "Excitation of surface wave on circular-loop array," *IEEE Trans. Antennas Propagat.*, **AP-19**, 433–435, 1971.

Yang, C. C. and K. C. Yeh, "The behavior of the backscattered power from an intensely turbulent ionosphere," *Radio Sci.*, **20**, 319–324, 1985.

Yang, Y. E., J. A. Kong, and Q. Gu, "Time domain perturbational analysis of nonuniformly coupled transmission lines," *IEEE Trans. Microwave Theory Tech.*, **MTT-33**, 1120–1130, 1985.

Yariv, A. and P. Yeh, *Optical Waves in Crystals*, Wiley-Interscience, New York, 1984.

Yeh, C. and K. F. Casey, "Reflection and transmission of electromagnetic waves by a moving dielectric slab," *Phys. Rev.*, **144**, 665–669, 1966.

Yeh, C., "Propagation along moving dielectric waveguides," *J. Opt. Soc. Am.*, **58**, 767–770, 1968.

Yeh, K. C., "Mutual coherence functions and intensities of backscattered signals in a turbulent medium," *Radio Sci.*, **18**, 159–165, 1983.

Yeh, K. C., "Second-order Faraday rotation formulas", *J. Geophys. Res.*, **65**, 2548–2550, 1960.

Yeh, K. C. and C. H. Liu, *Theory of Ionospheric Waves*, Academic Press, New York, 1972.

Zernike, F. and J. E. Midwinter, *Applied Nonlinear Optics*, Wiley-Interscience, New York, 1973.

Ziman, J. M., *Elements of Advanced Quantum Theory*, Cambridge University Press, London, 1969.

Zuniga, M. and J. A. Kong, "Modified radiative transfer theory for a two-layer random medium," *J. Appl. Phys.*, **51**, 5228–5244, 1980.